

STIC-Biotech/ChemLib

85802

From: Jiang, Dong  
Sent: Friday, January 31, 2003 7:16 PM  
To: STIC-Biotech/ChemLib  
Subject: 09/333,159

RECEIVED

FEB -3 2003

STIC-BIOTECH/ChemLib  
(STIC)

Please search SEQ ID NO: 47

Please send results on paper to Dong Jiang in 10D-08 (mail stop CM1-10D19).  
Thank you very much.

Dong Jiang (78243)  
703-305-1345  
U.S. Patent and Trademark Office  
Art Unit 1646  
dong.jiang@uspto.gov  
CM1-10D08  
Mail stop: CM1-10D19

Edward Hart  
Technical Info. Specialist  
STIC/Biotech  
CMI 6B02 Tel: 305-9203

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: 2/3/03  
Date Completed: 2/3/03  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH:

NA Sequences: \_\_\_\_\_  
AA Sequences: 1 \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)

STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: 02 \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

Pending Nucleic Acid and/or Pending Amino Acid database searches now generate two sets of results. These databases were split into two parts to reduce the time needed to update the databases daily. The split freed up more machine time for processing searches.

Searches run against the Nucleic Acid Pending database produce two sets of results, with the extensions, **.rnpm** and **.rnpn**

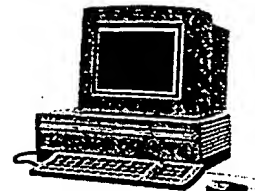
Searches run against the Amino Acid Pending database produce two sets of results, with the extensions, **.rapm** and **.rapn**

***The Pending database search results should not be left in the case because they contain data that is confidential.***

# BioTech-Chem Library

## Search Results

### Feedback Form (Optional)



Scientific & Technical Information C

The search results generated for your recent request are attached. If you have any questions or comments (compliments or complaints) about the scope or the results of the search, please contact *the BioTech-Chem searcher* who conducted the search *or contact*:

Mary Hale, Supervisor, 308-4:  
CM-1 Room 1E01

---

#### *Voluntary Results Feedback Form*

➤ *I am an examiner in Workgroup:* (Example: 1610)

➤ *Relevant prior art found, search results used as follows:*

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

*Types of relevant prior art found:*

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature  
(journal articles, conference proceedings, new product announcements etc.)

➤ *Relevant prior art not found:*

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Search results were not useful in determining patentability or understanding the invention.

**Other Comments:**

---

Drop off completed forms at the Circulation Desk CM-1, or send to Mary Hale, CM1-1E01 or [mary.hale@uspto.gov](mailto:mary.hale@uspto.gov)

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: February 3, 2003, 13:48:09; Search time 18 seconds  
(without alignments)  
691.439 Million cell updates/sec

Title: US-09-333-159-47  
Perfect score: 2247  
Sequence: 1 MLETLRQWIVSHRMEMWL.....IHLAQEETNLGRCRAVL 423

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: Issued\_Patents\_AA:  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/6C.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/Backfill1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1282	57.1	392	4	US-09-820-001-4 Sequence 4, Appl 1
2	1161	51.7	377	1	US-08-227-108-17 Sequence 17, Appl 1
3	1161	51.7	377	2	US-09-073-674-17 Sequence 17, Appl 1
4	1150.5	51.2	379	1	US-08-227-108-18 Sequence 18, Appl 1
5	1150.5	51.2	379	2	US-09-073-674-18 Sequence 18, Appl 1
6	1141.5	50.8	379	1	US-08-227-108-3 Sequence 3, Appl 1
7	1141.5	50.8	379	2	US-09-073-674-3 Sequence 3, Appl 1
8	1141.5	50.8	380	1	US-08-227-108-5 Sequence 5, Appl 1
9	1141.5	50.8	380	2	US-09-073-674-5 Sequence 5, Appl 1
10	1127	50.2	378	4	US-09-186-489-2 Sequence 2, Appl 1
11	1098.5	48.9	380	1	US-08-227-108-16 Sequence 16, Appl 1
12	1098.5	48.9	380	2	US-09-073-674-16 Sequence 16, Appl 1
13	1031	45.9	395	4	US-09-820-001-2 Sequence 2, Appl 1
14	105	4.7	430	1	US-08-484-105-22 Sequence 22, Appl 1
15	105	4.7	430	2	US-08-484-106-22 Sequence 22, Appl 1
16	102	4.5	436	1	US-08-602-359A-34 Sequence 34, Appl 1
17	98.5	4.4	388	1	US-08-232-519-2 Sequence 2, Appl 1
18	98.5	4.4	388	2	US-08-456-956-2 Sequence 2, Appl 1
19	91	4.0	298	4	US-09-355-166-6 Sequence 6, Appl 1
20	90	4.0	277	4	US-09-424-349A-6 Sequence 6, Appl 1
21	89.5	4.0	624	2	US-08-756-317-9 Sequence 9, Appl 1
22	89	4.0	494	4	US-09-134-001C-4475 Sequence 4475, Appl 1
23	89	4.0	1674	2	US-08-968-542C-12 Sequence 12, Appl 1
24	88.5	3.9	869	2	US-08-483-101-15 Sequence 15, Appl 1
25	84.5	3.8	338	2	US-08-602-359A-40 Sequence 40, Appl 1
26	84.5	3.8	935	1	US-07-707-367-2 Sequence 2, Appl 1
27	84	3.7	370	4	US-09-222-938A-64 Sequence 64, Appl 1

28	84	3.7	719	4	US-09-386-607-2 Sequence 2, Appl 1
29	83.5	3.7	341	4	US-09-412-600B-4 Sequence 4, Appl 1
30	82.5	3.7	485	2	US-08-446-803-1 Sequence 1, Appl 1
31	82.5	3.7	485	2	US-08-861-837-1 Sequence 1, Appl 1
32	82.5	3.7	485	2	US-08-600-908A-12 Sequence 12, Appl 1
33	82.5	3.7	485	3	US-08-683-838A-12 Sequence 12, Appl 1
34	82.5	3.7	485	3	US-08-600-656-1 Sequence 1, Appl 1
35	82.5	3.7	485	4	US-09-170-670-1 Sequence 1, Appl 1
36	82.5	3.7	485	4	US-09-170-670-7 Sequence 7, Appl 1
37	82.5	3.7	485	4	US-09-193-068-1 Sequence 1, Appl 1
38	82.5	3.7	485	4	US-09-193-068-7 Sequence 7, Appl 1
39	82.5	3.7	485	4	US-09-183-412-1 Sequence 1, Appl 1
40	82.5	3.7	485	4	US-09-183-412-7 Sequence 7, Appl 1
41	82.5	3.7	485	4	US-09-264-097-7 Sequence 7, Appl 1
42	82.5	3.7	485	4	US-09-354-191A-1 Sequence 1, Appl 1
43	82.5	3.7	485	4	US-09-291-023A-19 Sequence 19, Appl 1
44	82.5	3.7	485	4	US-09-290-734-1 Sequence 1, Appl 1
45	82.5	3.7	485	4	US-09-290-734-7 Sequence 7, Appl 1

## ALIGNMENTS

```
RESULT 1
US-09-820-001-4 Application US/09820001
; Sequence 4, Patent No. 6387680
; GENERAL INFORMATION:
; APPLICANT: MERKULOV, Genady et al
; TITLE OF INVENTION: ISOLATED HUMAN LIPASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN LIPASE PROTEINS, AND USES
; FILE REFERENCE: CLO01186
; CURRENT APPLICATION NUMBER: US/09/820,001
; CURRENT FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 392
; TYPE: PRT
; ORGANISM: Human
US-09-820-001-4
```

Query Match	57.1%	Score 1282;	DB 4;	Length 392;
Best Local Similarity	63.8%	Pred. No. 5.1e-127;		
Matches 234;	Conservative 53;	Mismatches 80;	Indels 0;	Gaps 0;
QY	40	AVDPEAFMNISETLHQGYPCBEYEVATEDGYLLSVNRIPRGVOPKKTGRPYVLLQHG	99	
DB	26	AVDPEETNMNVSEIISYWGPFSEYVETEDGYLLCLRIPHRKNSDKGPKPVFLQHG	85	
QY	100	LVGGAASNWSNLEPNNSLGFLLADAGFDVWNGNSRGNAKSKKHTLSIDQDEFAFSYDEM	159	
DB	86	LADSSNMVWVNLNNSLGFLLADAGFDVWNGNSRGNAKSKKHTLSVSDQEFNAFSYDEM	145	
QY	160	AREDLPAVIFLQKQGEKIVYVGSQGTTFMSTPELAOKIKMYALPIATVK	219	
DB	146	AKYDLPAVIFLQKQGEKIVYVGSQGTTFMSTPELAOKIKMYALPIATVK	205	
QY	220	HAKSPQTKELLPLDMIKGLFGKKEFLYQRFRLQVLYLVGQVYIDQISNMLLGGF	279	
DB	206	FCFSPMAKGLRDLPHDKLIDFGKKEFLYQRFRLQVLYLVGQVYIDQISNMLLGGF	265	
QY	280	NTNNMNAKRSVYAATHLATISQVNIIMHQAANSGLRAFDNCSSEKKNLEKQNPVPR	339	
DB	266	NEBNLNMSRDVYVTHSPACTSVQNMILHMSQAQFQFQAFDWCSSAKNFEHYQSYPT	325	
QY	340	YVRDVTVPYAMTGGODMNSPEDYKMLSEVTNLVYHNKIPFMAVDFIMGLDAPRM	399	
DB	326	YANKDLVYPAVAVSGHDMLADYDVNILLTQITNLVHFHSIFEMHLDIFMGLDAPRM	385	
QY	400	YNEIHL 406		

Db 249 KNLNVSREFDVYLGNPAGTSVQDFLHWAQLVRSCKFAFNWGPSQNM LHYNQKTPEYD 308

Qy	37	PTKAVDPEAPNISEIIQHOGYFCBEYEVATEDGYILSVNRPRLGLVQPKTGSRRPVLL	96
Db	7	PT--NPPNNNISQMLSYWGYPSEKYEVVTDGYILEVRNIPYKKNKSGNRQGRVPEL	63
Qy	97	QHGLVGASASWISNLPNNISLGFILADAGFDVWNGMNSRGNAWSRKHRTLSDIDQDEFWAFSY	156
Db	64	QHGLLASASWISNLPNNISLAFILADAGYGVWLGNSRGNTWSRNLYVSPDVSVEFWAFSF	123
Qy	157	DEMARFDLPVAVINILQKTQGEKIYVGYVSGQTMGFIAPSTWPELAQIKMYFALAPIA	216
Db	124	DEMAKYDLPATIDFIVAKTQGEKLHYVGHSGQTTIGFIAPSTPKLAERIKTFYALAPVA	183
Qy	217	TVKNAKSPGTRFLLPDMMTKGLGFKKEELYQTRFLRQLVYLCGOVILDQICSNTMLLL	276
Db	184	TVTKTSLVNLKLRIPPTMKIIFGDKIYFPHNFFDQFLATQVCSRETUNVICSNALFTI	243
Qy	277	GGFNTNNMNSRASVYAAHTLAGTSVQNLILHWSQVANSGBELADFDMGSETKNLEKNQPT	336

Query Match	51.2%	Score 1150.5;	DB 2;	Length 379;
Best Local Similarity	56.8%	Pred. NO. 3.9e-113;		
Matches 213:	Conservative	65;	Mismatches 94;	Indels 3;
				Gaps 1;

Qy	37	PTKAVDPAPFANNISEITQHQQYPCOEYEVEATEDGYILSVNRIPRGLVQPKKTSRGPVLL	96
Db	7	PT--NPEVNNNISOMISYWGYPSEKYEVTEDGYILEVNRIPYKKNSGNRGQRPVFL	63
Qy	97	QHGLVGGASNIWLNLPNNLSGLFIADAGFDVWNGNSRGNASHKHKLTLSIDODEFAFSY	150
Db	64	QHGLLASASNIWLNLPNNLSLAFILADAGGYVWLGNSRGNTWSRNLNYSPODSVEFAFSE	123
Qy	157	DEMARFDLPVINFILQKTGGEKYVYVCYSOGTTWGFTAFSTMPELQAKIKMYFALAPIA	216
Db	124	DEMAKYDLPATIDFVKETGGEKLVYVGHSGOTTIGFTAFSTNPKLAERIKTFYALAPVA	183
Qy	217	TVVHKASGTFKLLLPDMNMIKGLSGKKFELQTFRLQRLVLYLGCQVILDOICSNIMLLL	270

Db 184 TVKYTKSLVKNLRFIPPTMFKIIFGDKIYFPHNFDFQFLATQVCSRETLNVIGSNALFII 243  
QY 277 GGFNTNMNMSRASYAAHTLAGTSVQNILHWSQAVNSGELRAFDMGSETKNLEKCNQPT 336  
Db 244 CGFDSANLNSRLDVYSHNPAGTSVQNILHWTQAVKSGNFQAFNGSPAQNVYHFNQPT 303  
QY 337 PVRYRDMTVPMTAMTGGODMLSNPDKMLLSEVTNLIYHKNIPWHAHVDIFWGLDAP 396  
Db 304 PPYVNTAMNVPJAVNSGGNDWLDQDDVLLPLKLSNLIIYHKEILLPYNHLDFIWMANAP 363  
QY 397 HRMYNEIHLMOQEE 411  
Db 364 QEVYNEIISMAKDK 378

## RESULT 6

US-08-227-108-3  
; Sequence 3, Application US/08227108  
; Patent No. 5807726  
; GENERAL INFORMATION:  
; APPLICANT: Blanchard, Claire  
; APPLICANT: Benicourt, Claude  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/227,108  
; FILING DATE: 03-APR-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fanucci, Allan A.  
; REGISTRATION NUMBER: 30,256  
; REFERENCE/DOCKET NUMBER: 7620-033  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212 790-9090  
; TELEFAX: 212 869-8864/9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 379 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-227-108-3

Query Match 50.8%; Score 1141.5; DB 1; Length 379;  
Best Local Similarity 56.5%; Pred. No. 3.4e-112;  
Matches 210; Conservative 67; Mismatches 90; Indels 5; Gaps 3;

QY 37 PTKAVDPEAFMNISEIIHQGYPCEEYEVATEDGYILSVNRIPRGLVQPKKTSRPPVLL 96  
Db 7 PT---NPEVTMNIISQIMTYGYPAAEYEVVTEGDYILGIDRIPYGRKNSENIGRRPVAFL 63  
QY 97 QHGLVGASWISNLPNNSLGLFADAGFDVWNGSGNNAWSRKHHTLSIDQDEFWAFSY 156  
Db 64 QHGLASATWISNLPNNSLAFILADAGYDVLGNSRGNTWARRNLVYSDSPDEFWAFSF 123  
QY 157 DEMARDELPAVINFILOKTOGEKIYYGVYSGQTTMGFIATSTPELAQKIKMYFALAPIA 216  
Db 124 DEMAKYDLPATIOFILKTKQDKLHVYGHSGQTTIGFIATSTPNKAKRIKTFYALAPVA 183

QY 217 TVKHAKSPGTFKLLPDMIMKGLFGKKEFLYQTRFLRQ-LVIVLCGOVILDOICSNIMLL 275  
Db 184 TVKYTETLLNKLMLVPSFLKLFIGNKIF-YPHFFDOFLATEVCSRETYDLLCSNALFI 242  
QY 276 LGGFNTNMNMSRASYAAHTLAGTSVQNILHWSQAVNSGELRAFDMGSETKNLEKCNQPT 335  
Db 243 ICGFDTMNLNLSRLDVYSHNPAGTSVQNILHWSQAVKSGNFQAFNGSPQVNNMHHQHS 302  
QY 336 TPVRYRDMTVPMTAMTGGODMLSNPDKMLLSEVTNLIYHKNIPWHAHVDIFWGLDAP 396  
Db 304 PPYVNTAMNVPJAVNSGGNDWLDQDDVLLPLKLSNLIIYHKEILLPYNHLDFIWMANAP 363  
QY 397 HRMYNEIHLMOQEE 411  
Db 364 QEVYNEIISMAKDK 378

## RESULT 7

US-09-073-674-3  
; Sequence 3, Application US/09073674  
; Patent No. 5998189  
; GENERAL INFORMATION:  
; APPLICANT: Blanchard, Claire  
; APPLICANT: Benicourt, Claude  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Warner-Lambert Company  
; STREET: 2800 Plymouth Road  
; CITY: Ann Arbor  
; STATE: Michigan  
; COUNTRY: U.S.A.  
; ZIP: 48105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/073,674  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Crissey, Todd M.  
; REGISTRATION NUMBER: 37,807  
; REFERENCE/DOCKET NUMBER: 5072-D1-66-TWC  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 734 622-7530  
; TELEFAX: 734 622-1553  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 379 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-073-674-3

Query Match 50.8%; Score 1141.5; DB 2; Length 379;  
Best Local Similarity 56.5%; Pred. No. 3.4e-112;  
Matches 210; Conservative 67; Mismatches 90; Indels 5; Gaps 3;

QY 37 PTKAVDPEAFMNISEIIHQGYPCEEYEVATEDGYILSVNRIPRGLVQPKKTSRPPVLL 96  
Db 7 PT---NPEVTMNIISQIMTYGYPAAEYEVVTEGDYILGIDRIPYGRKNSENIGRRPVAFL 63  
QY 97 QHGLVGASWISNLPNNSLGLFADAGFDVWNGSGNNAWSRKHHTLSIDQDEFWAFSY 156  
Db 64 QHGLASATWISNLPNNSLAFILADAGYDVLGNSRGNTWARRNLVYSDSPDEFWAFSF 123  
QY 157 DEMARDELPAVINFILOKTOGEKIYYGVYSGQTTMGFIATSTPELAQKIKMYFALAPIA 216

Db 124 DEMAKYDLPATIDFLLKKTGQDLHYVHSGQTTIGFIAFSTNPKLAKRIKTFYALAPVA 183  
QY 217 TVKHAKSPGKFKLLPDMITKGLFGKKEFLYQTRFLRQ-LVIVLCGOVILDOICSNIMLL 275  
Db 184 TVKYTELLNKLMLVPSFLFKLIFGNKIF-YPHHFFDQFLATEVCSRETVDLLCSNALFI 242  
QY 276 LGGFNTNNMNSRASVYAAHTLAGTSVQNLHWSQAVNSGELRAFDWGSRTKLNKRCNP 335  
Db 243 ICGFDTMNLNMSRLDYLSHNPAQTSVQNVLHWSQAVKSGKFOAFDWSGPVQNMHHYQS 302  
QY 336 TPVRYVRDMTVPMTAMTGGQDLSNPEDVKMLLSEVTNLIYKHNIPWAHVDFIWLGLDA 395  
Db 303 MPYYNLTDHVPVIAVWNGNDLLADPHDVLKLPNLIYHRKIPPNYHLDFIWAMDA 362  
QY 396 PHRYNEIHLM 407  
Db 363 PQAVYNEIVSM 374

## RESULT 8

US-08-227-108-5  
; Sequence 5, Application US/08227108  
; Patent No. 5807726  
; GENERAL INFORMATION:  
; APPLICANT: Blanchard, Claire  
; APPLICANT: Benicourt, Claude  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/227,108  
; FILING DATE: 03-APR-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Panucci, Allen A.  
; REGISTRATION NUMBER: 30,256  
; REFERENCE/DOCKET NUMBER: 7620-033  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212 790-9090  
; TELEFAX: 212 869-8864/9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 380 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-227-108-5

Query Match 50.8%; Score 1141.5; DB 1; Length 380;  
Best Local Similarity 56.5%; Pred. No. 3.5e-112;  
Matches 210; Conservative 67; Mismatches 90; Indels 5; Gaps 3;  
QY 37 PTKAVDPEAFNISEIIOHOGYPCPEEYEVATEDGYILSVNRIPRGLVQPKTGRSPVLL 96  
Db 8 PT---NPEVTWNISQITMYGPAEYEVVTDGYILGIDRIPYGRKNSENIGRRPVAFL 64  
QY 97 QHGLVGASWISNLPNNLSGLFILDAGFDVWNGNSRGNWSRKHKTLSIDQDDEFWAFSY 156  
Db 65 QHGLLASATWNISNLPNNLSGLFILDAGFDVWNGNSRGTWARRNLYSPDSVBEWAFSE 124

QY 157 DEMAREFDPAVINFILOKTGOEKIYVYVSGQTTMGFIAFSTMPELAQKIKMYFALAPIA 216  
Db 125 DENAKYDLPATIDFLLKKTGQDLHYVHSGQTTIGFIAFSTNPKLAKRIKTFYALAPVA 184  
QY 217 TVKHAKSPGKFKLLPDMITKGLFGKKEFLYQTRFLRQ-LVIVLCGOVILDOICSNIMLL 275  
Db 185 TVKYTELLNKLMLVPSFLFKLIFGNKIF-YPHHFFDQFLATEVCSRETVDLLCSNALFI 243  
QY 276 LGGFNTNNMNSRASVYAAHTLAGTSVQNLHWSQAVNSGELRAFDWGSRTKLNKRCNP 335  
Db 244 ICGFDTMNLNMSRLDYLSHNPAQTSVQNVLHWSQAVKSGKFOAFDWSGPVQNMHHYQS 303  
QY 336 TPVRYVRDMTVPMTAMTGGQDLSNPEDVKMLLSEVTNLIYKHNIPWAHVDFIWLGLDA 395  
Db 304 MPYYNLTDHVPVIAVWNGNDLLADPHDVLKLPNLIYHRKIPPNYHLDFIWAMDA 363  
QY 396 PHRYNEIHLM 407  
Db 364 PQAVYNEIVSM 375

## RESULT 9

US-09-073-674-5  
; Sequence 5, Application US/09073674  
; Patent No. 5998189  
; GENERAL INFORMATION:  
; APPLICANT: Blanchard, Claire  
; APPLICANT: Benicourt, Claude  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Warner-Lambert Company  
; STREET: 2800 Plymouth Road  
; CITY: Ann Arbor  
; STATE: Michigan  
; COUNTRY: U.S.A.  
; ZIP: 48105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/073,674  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Crissey, Todd M.  
; REGISTRATION NUMBER: 37,807  
; REFERENCE/DOCKET NUMBER: 5072-D1-66-TMC  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 734 622-7530  
; TELEFAX: 734 622-1553  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 380 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-073-674-5

Query Match 50.8%; Score 1141.5; DB 2; Length 380;  
Best Local Similarity 56.5%; Pred. No. 3.5e-112;  
Matches 210; Conservative 67; Mismatches 90; Indels 5; Gaps 3;  
QY 37 PTKAVDPEAFNISEIIOHOGYPCPEEYEVATEDGYILSVNRIPRGLVQPKTGRSPVLL 96  
Db 8 PT---NPEVTWNISQITMYGPAEYEVVTDGYILGIDRIPYGRKNSENIGRRPVAFL 64  
QY 97 QHGLVGASWISNLPNNLSGLFILDAGFDVWNGNSRGNWSRKHKTLSIDQDDEFWAFSY 156  
Db 65 QHGLLASATWNISNLPNNLSGLFILDAGFDVWNGNSRGTWARRNLYSPDSVBEWAFSE 124



Db 65 QHGLASATNWSINLNNLAFILADAGYDVLGNSRGNTWARRNLYYSPDSVEFWAFSF 124  
QY 157 DEMARFDLPVNFILQKTCQKIIYVGYSGQTGTFIAFSMPELAOKIKMYFALAPTA 216  
Db 125 DEMAYDLPATIDFLIKTKGDKLHVHSGQTGTFIAFSMPKLAKRIKTFYALAPYA 184  
QY 217 TVKHAKSPCTKLLPDMNKGKFKKELYOTRELQ-LVIYLCGOVILDOICSNIMLL 275  
Db 185 TVKYTETLLNKLMLVPSFLFKLFGNKF-YPHHFDQFLATEVCSRETVDDLCSNALFI 243  
QY 276 LGGFNTNMMNSRASVYAAHTLAGTSVONILHWSQAVNSGELRAFDMGSETKNLEKCNQ 335  
Db 244 ICGFOTMLNMSRLDVLHSHNPAGTSVQVNLHWSQAVKSGKFAEDMGSPVQNMHHQS 303  
QY 336 TPVRYVRDMTPTAMWTCGQDLNPNEDVKMLLSEVTNLIYHKNIPENAHVDFIWLDA 395  
Db 304 MPYYNLTDMHVPVIAVWNGNDLADPHDVLKLLSLNLIYHKKIPPNYHLDIFWAMDA 363  
QY 396 PHRYMNEIILM 407  
Db 364 POAVNEIVSMM 375  
RESULT 10  
US-09-186-489-2  
; Sequence 2, Application US/09186489  
; Patent No. 6375947  
; GENERAL INFORMATION:  
; APPLICANT: Bolen, Paul L.  
; APPLICANT: Cihak, Paul L.  
; APPLICANT: Scharpf Jr., Lewis G.  
; TITLE OF INVENTION: Purified Recombinant Kid Pregastric Esterase, and  
; TITLE OF INVENTION: Processes for its Production and Use  
; FILE REFERENCE: 5499/3  
; CURRENT APPLICATION NUMBER: US/09/186,489  
; CURRENT FILING DATE: 1998-11-05  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 378  
; TYPE: PRT  
; ORGANISM: Kid (Goat)  
US-09-186-489-2  
Query Match 50.2%; Score 1127; DB 4; Length 378;  
Best Local Similarity 55.1%; Pred. No. 1.2e-110;  
Matches 205; Conservative 66; Mismatches 101; Indels 0; Gaps 0;  
QY 40 AYDPEAFMISEIIOHGYPCPEYEVATEDGYILSVNRIPIRGVLQPKKTSRPPVLLQHG 99  
Db 6 ANPEASMVWSQISWGPSEMHKVTADGYILQVYRIPHGKNDANHLQRPVFLQHG 65  
QY 100 LVGGASNWSINLNNLPGFLADAGYDVLGNSRGNTWARRNLYYSPDSVEFWAFSD 159  
Db 66 LLASATNWSINLNNLPGFLADAGYDVLGNSRGNTWARRNLYYSPDSVEFWAFSD 125  
QY 160 ARFDLPVNFILQKTCQKIIYVGYSGQTGTFIAFSMPELAOKIKMYFALAPATVK 219  
Db 126 AEYDLPSTIDFLIKTKGDKLHVHSGQTGTFIAFSMPKLAKRIKTFYALAPATVK 185  
QY 220 HAKSPCTKLLPDMNKGKFKKELYOTRELQ-LVIYLCGOVILDOICSNIMLLGGF 279  
Db 186 HTQSLFNKLALIPHEFLKIFGNKMYPHNPFQFLGVCSRETLDVLCKNALFAITGA 245  
QY 280 NTNMMNSRASVYAAHTLAGTSVONILHWSQAVNSGELRAFDMGSETKNLEKCNQPT 339  
Db 246 DKNFNMSRLDVLHSHNPAGTSVQVNLHWSQAVKSGKFAEDMGSPVQNMHHQS 305  
QY 340 YVRDVTPTAMWTCGQDLNPNEDVKMLLSEVTNLIYHKNIPENAHVDFIWLDA 399  
Db 306 YNLTMVPIAVWSAGQDLNPNEDVKMLLSEVTNLIYHKNIPENAHVDFIWLDA 363  
QY 400 YNEIILMQEE 411

Db 366 YNEIILMAKOK 377  
RESULT 11  
US-08-227-108-16  
; Sequence 16, Application US/08227108  
; Patent No. 5807726  
; GENERAL INFORMATION:  
; APPLICANT: Blanchard, Claire  
; APPLICANT: Benicourt, Claude  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/227,108  
; FILING DATE: 03-APR-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fanucci, Allan A.  
; REGISTRATION NUMBER: 30,256  
; REFERENCE/DOCKET NUMBER: 7620-033  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212 790-9090  
; TELEFAX: 212 869-8864/9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 380 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-227-108-16  
Query Match 48.9%; Score 1098.5; DB 1; Length 380;  
Best Local Similarity 54.6%; Pred. No. 1.2e-107;  
Matches 202; Conservative 69; Mismatches 98; Indels 1; Gaps 1;  
QY 43 PEAFMISEIIOHGYPCPEYEVATEDGYILSVNRIPIRGVLQPKKTSRPPVLLQHLVG 102  
Db 10 PEVTMISOMITWGPSEMHKVTADGYILQVYRIPHGKNDANHLQRPVFLQHLA 69  
QY 103 GASWISNLSNPLSGFLADAGYDVLGNSRGNTWARRNLYYSPDSVEFWAFSD 161  
Db 70 SATNWSINLNNLPGFLADAGYDVLGNSRGNTWARRNLYYSPDSVEFWAFSD 129  
QY 162 FDLPAVINFLQKTCQKIIYVGYSGQTGTFIAFSMPELAOKIKMYFALAPATVK 221  
Db 130 YDLPATIDFIVKTKGDKLHVHSGQTGTFIAFSMPKLAKRIKTFYALAPATVK 189  
QY 222 KSPGTKFLLLPDMNKGKFKKELYOTRELQ-LVIYLCGOVILDOICSNIMLLGGFNT 281  
Db 190 KSLINKLRFVPSLKFIFGDKIFYPHNFDDQFLATEVCSREMLNLCNALFIICGDS 249  
QY 282 NNMMNSRASVYAAHTLAGTSVONILHWSQAVNSGELRAFDMGSETKNLEKCNQPT 341  
Db 250 KNFTSLDVLHSHNPAGTSVQVNLHWSQAVKSGKFAEDMGSPVQNMHHQS 309  
QY 342 VRDVTPTAMWTCGQDLNPNEDVKMLLSEVTNLIYHKNIPENAHVDFIWLDA 401

Db 310 VTAMNPIAVNNGKDLLADPODVGLLPLKLPNLIYHKEIPFYNHLDPIWADAPQEVYN 369

Qy 402 EIIHLMQOEE 411

Db 370 DIVSMISEDK 379

## RESULT 12

US-09-073-674-16  
; Sequence 16, Application US/09073674

; Patent No. 5998189

; GENERAL INFORMATION:

; APPLICANT: Blanchard, Claire

; APPLICANT: Benicourt, Claude

; APPLICANT: Junien, Jean-Louis

; TITLE OF INVENTION: Recombinant Dog Gastric Lipase

; NUMBER OF SEQUENCES: 21

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Warner-Lambert Company

; STREET: 2800 Plymouth Road

; CITY: Ann Arbor

; STATE: Michigan

; COUNTRY: U.S.A.

; ZIP: 48105

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/073.674

; FILING DATE:

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Crissey, Todd M.

; REGISTRATION NUMBER: 37,807

; REFERENCE/DOCKET NUMBER: 5072-D1-66-TMC

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 734 622-7530

; TELEFAX: 734 622-1553

; TELEX:

; INFORMATION FOR SEQ ID NO: 16:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 380 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-09-073-674-16

Query Match 48.9%; Score 1098.5; DB 2; Length 380;

Best Local Similarity 54.6%; Pred. No. 1.2e-107;

Matches 202; Conservative 69; Mismatches 99; Indels 1; Gaps 1;

Qy 43 PFAFNISIIHQOGYPCPEYEVATEDGYLSVNRIPRGLVQPKTGRPVVLLQRLGV 102

Db 10 PEVTNISMOMITYGYNBEEYEWVTEGYLEVNRIPYKKNKSGTGQRPVVLQGLLA 69

Qy 103 GASNWSISLNNLSGLFILLADAGFDVWNGSRGNASRKHKTLSIDQDEFW-AFSDEMAR 161

Db 70 SATNWSISLNNLSGLFILLADAGFDVWNGSRGNATWARRNLYYSPDSVEFWAFSDEMAK 129

Qy 162 FDLPAVINFLQKTGOEKIYVYVYSGTTMGTFATFAPTELAOKIKMYFALAPIATVKHA 221

Db 130 YDLPATIDFIVKTKQKQLHYVGHSGQTIGTGFATFSTWPSLAKRIKTFYALAPVATVIT 189

Qy 222 KSPGTFKLLPDMKIMKGLFGKKEFLYQTRFLRLQVLYLQGVILQDICSNMILLGGFT 281

Db 190 KSLINKLRFVPOSFLFKFIDGKIFYPHNFDDQFLATEVCSREMLNLLCSNALFIICGDS 249

Qy 282 NNNMSRASVYAAHTLAGTSVQNLHWSQAVNSGELRAFWDGSETKNEKCNQPTPVYR 341

Db 250 KNFTSRLDVLVYLSHPAGTSVQNMFWHTQAVKSGKFOAYDWGSPVQNRMRHYDQSQPPYIN 309

Qy 342 VRDMTVPMTAGTGGDWLSNPEDVKMLLSEVTNLIYHKNIPWAHVDFTWGLDAPHMYN 401

Db 310 VTAMNPIAVNNGGKDLLADPODVGLLPLKLPNLIYHKEIPFYNHLDPIWADAPQEVYN 369

Qy 402 EIIHLMQOEE 411

Db 370 DIVSMISEDK 379

## RESULT 13

US-09-820-001-2

; Sequence 2, Application US/09820001

; Patent No. 6387680

; GENERAL INFORMATION:

; APPLICANT: MERKULOV, Gennady et al

; TITLE OF INVENTION: ISOLATED HUMAN LIPASE PROTEINS, NUCLEIC

; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN LIPASE PROTEINS, AND USES

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: CLO01186

; CURRENT APPLICATION NUMBER: US/09/820.001

; CURRENT FILING DATE: 2001-03-29

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 395

; TYPE: PRT

; ORGANISM: Human

; US-09-820-001-2

Query Match 45.9%; Score 1031; DB 4; Length 395;

Best Local Similarity 51.0%; Pred. No. 1.8e-100;

Matches 199; Conservative 61; Mismatches 124; Indels 6; Gaps 4;

Qy 17 MWLLILVAYMPQNVNS---VHMPTKAVDPPEAFMNISEIIHQGIPCEYEVATEDGYIL 73

Db 2 MWLLITTTCLICGTLNAGGFLDLENE-VNPEYVMMNTSEIIYNGVPSEYEVATEDGYIL 60

Qy 74 SVNRIPLGLVOPKKTGRPVVLLQHLGVGASNSLNPLNNSIGFTLADAGFDVWNGSR 133

Db 61 LVNRIPLGTRHARSTGPRPVVYMHAFADNAYLENYANGSLGFLAAGFDVWNGSR 120

Qy 134 GNWSRKHKTLSIDODEFWAFSYDEMAREDLPAVINFLQKTGOEKIYVYVYSGTTMGF 193

Db 121 GNEWSRRHKTLSIDODEFWAFSYDEMAREDLPAVINFLQKTGOEKIYVYVYSGTTMGF 180

Qy 194 TAFSTMPELAQIKMYFALAPIATVKHAKSPGTFKLLPDMKIMKGLFGKKEFLYQTRFLR 253

Db 181 VAFSTMPELAQIKMYFALAPIATVKHAKSPGTFKLLPDMKIMKGLFGKKEFLYQTRFLR 240

Qy 254 QLVVYLCGOVILDOICSNIMLLGGFTNMMNMSRASVYAAHTLAGTSVQNLHWSQAVN 313

Db 241 TASTKICNNKILWLICSEFMSLWAGSNKKMMNOSRMDVYMSHAPTSSVHNILHKLQVH 300

Qy 314 SGEELRAFDMGSETKNEKCNQPTPVYRVRDMTQVPTAMTGGODMWSNPEDVKMLLSEVT 373

Db 301 SDEFRAYDWDGNDADNNKHYNQSHPPYIDLTAKKPYTAIWAGGHVLTGTQDVARILPQIK 360

Qy 374 NL-IYHKNIPWAHY-DFIWLGLDAPHMYN 401

Db 361 SLSLSLSLPEWEPTDFVWGLDAPQRMFS 390

## RESULT 14

US-08-484-105-22

; Sequence 22, Application US/08484105

; Patent No. 5589341

; GENERAL INFORMATION:

; APPLICANT: STILLMAN, Bruce

; APPLICANT: BELL, Stephen P

; APPLICANT: KOBAYASHI, Ryuji

; APPLICANT: RINE, Jasper

; APPLICANT: FOSS, Margit

APPLICANT: FOSS, Margit  
 APPLICANT: MCNALLY, Francis J  
 APPLICANT: LAURENSEN, Patricia  
 APPLICANT: HERSKOWITZ, Ira  
 APPLICANT: LI, Joachim J  
 APPLICANT: GAVIN, Kimberly  
 TITLE OF INVENTION: ORIGIN OF REPLICATION COMPLEX GENES  
 NUMBER OF SEQUENCES: 24  
 CORRESPONDENCE ADDRESS:  
 ADDRESSSEE: FLEHR, HOOBACH, TEST, ALBRITTON & HERBERT  
 STREET: 4 Embarcadero Center, Suite 3400  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94111-4187  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/484.106  
 FILING DATE:  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Osman Ph.D., Richard Aton  
 REGISTRATION NUMBER: 36,627  
 REFERENCE/DOCKET NUMBER: A-59032/DJB/RAO  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 494-8700  
 TELEFAX: (415) 494-8771  
 TELEX: 910 277299  
 INFORMATION FOR SEQ ID NO: 22:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 430 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-484-106-22

```

Query Match      4.7%; Score 105; DB 1; Length 430;
Best Local Similarity 21.9%; Pred. No. 0.016;
Matches 61; Conservative 42; Mismatches. 87; Indels 88; Gaps 13;

y y      85 PKTGRPPVLLQHLGLVGASWISLNPNN-----SLGF-----ILADAGEDVW--MGN 131
      ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
b      25 PEKEGR-----QKKTNGKENASRLQSLNEEDLEQLGFDEFTVSMQASAIENFMQK 78
      ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
y      132 S----RGNAWSRKHK-----TLSIDDDFWAFSYDEMARFDLPVINFILQKTGOEKIY 181
      ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
b      79 SAGERNNNAKRRGRAGNGNTEEBEEDDEISNAITDFTKCDLPLGLRNITTKNDNTEPEK 138
      ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
      182 YVGYSOQTNGGTFATSTMPELAQIKWYFALAPIATVKHAKSPGTFKLLLPDMWIKGLFG 241
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
b      139 RLEHLADNDFG-----KWKLYLAAG-----FWILLHGVS 168
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
      242 KKEFLYQTRFLRQLVYILCGQVILDCISNIMLLGGFNIN---NNMNSRASVYAAHTLA 298
      ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
b      169 KRDLV--TEENELSDTYMRVDARDGLNKKVLLGAIENNMKLNCKVRG----- 217
      ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
y      299 GTSVQNILHWSQAV----NSGEL-----RAPDWGSE 325
      ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
b      218 ----QSTISWARRRKMNSQOOLIIDNIEAPDWRSD 251
      ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 15  
US - 08-484-106-22  
; Section 22, Application US/08484106  
; Patent No. 5614618  
; GENERAL INFORMATION:  
; APPLICANT: STILLMAN, Bruce  
; APPLICANT: BELL, Stephen P  
; APPLICANT: KOBAYASHI, Ryuji  
; APPLICANT: RINE, Jasper

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:44:29 ; Search time 40 Seconds  
(without alignments)  
1409.125 Million cell updates/sec

Title: US-09-333-159-47

Perfect score: 2247

Sequence: 1 MLETLSRQWIVSHRMEMWLL.....IHLMQEETNLSQGRCEAVL 423

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Result No.	Score	Query Match	Length	ID	Description
1	2247	100.0	423	22 AAB66065	Human TANGO 294.
2	2174	96.8	409	22 AAE11931	Human CGI162 (or C5
3	2076	92.4	390	22 AAB66067	Human TANGO 294 ma
4	1991	88.6	398	23 AAU99164	Human lysosomal ac
5	1536.5	68.4	371	23 AAU99165	Human lysosomal ac
6	1289	57.4	399	22 AAB200783	Human shear stress
7	1289	57.4	399	22 AAB66061	Human lysosomal ac
8	1282	57.1	392	22 AAO18227	Human lysosomal ac
9	1191.5	53.0	449	22 AAU30498	Novel human secret
10	1166	51.9	398	14 AAR37302	RGL precursor. Or

#### SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

11	1163	51.8	221	22 AAB66068	Human TANGO 294 ex
12	1156.5	51.5	395	6 AAP50322	Rat lingual lipase
13	1143	50.9	395	23 AAE14744	Human triacylglycer
14	1143	50.9	399	23 AAU77496	Human lipid metabo
15	1141.5	50.8	379	15 AAR56870	Canine gastric lip
16	1141.5	50.8	379	17 AAU09382	Dog gastric lipase
17	1141.5	50.8	380	15 AAR56871	Canine gastric lip
18	1139	50.7	399	22 AAG67513	Amino acid sequenc
19	1135	50.6	397	23 AAE17308	Human lysosomal ac
20	1127	50.2	378	23 ABB76189	Kid goat pregestril
21	1119	49.8	398	7 AAP60724	Sequence of pregas
22	1119	49.8	398	7 AAP60658	Sequence of human
23	1119	49.8	398	17 AAU09383	Human gastric lipa
24	1119	49.8	398	22 AAB66086	Human lipase prote
25	1101	49.0	403	23 AAU77493	Human lipid metabo
26	1090	48.5	427	23 AAU98539	Human lysosomal ac
27	1031	45.9	395	23 AAO18226	Human lysosomal ac
28	977	43.5	365	23 AAE17307	Human lysosomal ac
29	907	40.4	731	23 AAE14746	Human triacylglycer
30	783	34.8	144	22 AAB66070	Human TANGO 294 cy
31	781	34.8	276	23 AAE14745	Human triacylglycer
32	777	34.6	289	23 AAU77494	Human lipid metabo
33	745.5	33.2	280	23 AAE14743	Human triacylglycer
34	724.5	32.2	656	22 ABG26839	Novel human diagne
35	691	30.8	233	22 AAB61608	Human protein HP03
36	663	29.5	394	22 ABB71702	Drosophila melanog
37	655.5	29.2	434	22 ABB62081	Drosophila melanog
38	639	28.4	456	22 ABB68957	Drosophila melanog
39	636	28.3	398	22 ABB63011	Drosophila melanog
40	627	27.9	311	23 AAU98540	Human lysosomal ac
41	588.5	26.2	838	22 ABB61967	Drosophila melanog
42	567.5	25.3	457	22 ABB63187	Drosophila melanog
43	559	24.9	399	22 ABB71442	Drosophila melanog
44	546	24.3	439	22 ABB63143	Drosophila melanog
45	535	23.8	416	22 ABB59328	Drosophila melanog

#### ALIGNMENTS

RESULT 1  
AAB66065  
ID AAB66065 standard; Protein: 423 AA.  
XX AC AAB66065;  
XX DT 30-MAR-2001 (first entry)  
XX DE Human TANGO 294.

XX KW TANGO protein; INTERCEPT protein; neurological disorder;  
XX KW central nervous system; focal brain disorder; bipolar affective disorder;  
XX KW global diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;  
XX KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;  
XX KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;  
XX KW neuropsychiatric; psychoactive substance use; anxiety.

OS Homo sapiens.

XX WO200077239-A2.

XX PD 21-DEC-2000.

XX PF 24-MAY-2000; 2000WO-US14858.

XX PR 14-JUN-1999; 99US-0333159.

XX PA (MILL-) MILLENNIUM PHARM INC.

XX PI McCarthy SA, Fraser CC, Sharp JD, Barnes TM;

XX DR WPI; 2001-032313/04.

XX DR N-PSDB; AAF45131, AAF45132.



|||||  
Db 61 VNRIPGLVOPKTCGRPVVLLQHLVGGASNNISNLPNNLSGLFADAGFDVWNGNSRG 120  
QY 135 NANSRKHKTLSIDQDEFWAFSYDEMAREFDLPVAVINFILQKTGOEKIYVYGYSQGTMGFI 194  
Db 121 NANSRKHKTLSIDQDEFWAFSYDEMAREFDLPVAVINFILQKTGOEKIYVYGYSQGTMGFI 180  
QY 195 AFSTMPPELAQIKMYFALAPVATVYKHAKEGTFKLLPDMNKGKKEFLYQTRFLRQ 254  
Db 181 AFSTMPPELAQIKMYFALAPVATVYKHAKEGTFKLLPDMNKGKKEFLYQTRFLRQ 240  
QY 255 LVLYLCCQVILQDQCSNIMLLGFTNNMNSRASVYAAHTLAGTSVQNLHWSQAVNS 314  
Db 241 LVLYLCCQVILQDQCSNIMLLGFTNNMNSRASVYAAHTLAGTSVQNLHWSQAVNS 300  
QY 315 GELRAFQWSETKLNKCNQPTPVYRVRDMTPTAMWTGGQDWSNPEDVKMLLSEVTN 374  
Db 301 GELRAFQWSETKLNKCNQPTPVYRVRDMTPTAMWTGGQDWSNPEDVKMLLSEVTN 360  
QY 375 LIYHKNIPPEAHVDFINGLQDAPHRMYNEIHLMQOQETNLSQGRCEAVL 423  
Db 361 LIYHKNIPPEAHVDFINGLQDAPHRMYNEIHLMQOQETNLSQGRCEAVL 409

RESULT 3  
AAB66067  
ID AAB66067 standard; Protein: 390 AA.  
XX AC AAB66067;  
XX 30-MAR-2001 (first entry)  
DE Human TANGO 294 mature protein.  
XX  
KW TANGO protein; INTERCEPT protein; neurological disorder;  
KW central nervous system; focal brain disorder; bipolar affective disorder;  
KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;  
KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;  
KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;  
KW neuropsychiatric; psychoactive substance use; anxiety.  
OS Homo sapiens.  
XX  
XX WO200077239-A2.  
XX  
XX 21-DEC-2000.  
XX  
XX 24-MAY-2000; 2000WO-US14858.  
XX  
XX 14-JUN-1999; 99US-0333159.  
XX (MILL-) MILLENNIUM PHARM INC.  
XX PA McCarthy SA, Fraser CC, Sharp JD, Barnes TM;  
XX PI  
XX WPI; 2001-032313/04.  
XX  
XX TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for  
PT screening assays and diagnostic assays and for the treatment of  
PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's  
PT disease -  
XX  
XX Claim 8; Pages 324-325; 359pp; English.  
XX  
XX The present invention relates to TANGO or INTERCEPT proteins and coding  
CC sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,  
CC AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding  
CC sequences are useful for the treatment of neurological disorders such as  
CC central nervous system (CNS) disorders, CNS-related disorders, focal  
CC brain disorders, global-diffuse cerebral disorders and other  
CC neurological and cerebrovascular disorders. The CNS disorders include  
CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic  
CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,

CC autonomic function disorders such as hypertension and sleep disorders,  
CC neuropsychiatric disorders, psychoactive substance use disorders,  
CC anxiety, and bipolar affective disorder.  
XX  
SQ Sequence 390 AA;  
Query Match 92.4%; Score 2076; DB 22; Length 390;  
Best Local Similarity 100.0%; Pred. No. 1.9e-202;  
Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 34 VHMPTKAVDPEAFNMISEIIHQHGYPCPEEVEVATEDGYILSVNRIPRGLVQPKKTGSRPV 93  
Db 1 VHMPTKAVDPEAFNMISEIIHQHGYPCPEEVEVATEDGYILSVNRIPRGLVQPKKTGSRPV 60  
QY 94 VLLQHLVGGASNNISNLPNNLSGLFADAGFDVWNGNSRGNASRKHKTLSIDQDEFWA 153  
Db 61 VLLQHLVGGASNNISNLPNNLSGLFADAGFDVWNGNSRGNASRKHKTLSIDQDEFWA 120  
QY 154 FSYDEMAREFDLPVAVINFILQKTGOEKIYVYGYSQGTMGFIATSTMPPELAQIKMYFALA 213  
Db 121 FSYDEMAREFDLPVAVINFILQKTGOEKIYVYGYSQGTMGFIATSTMPPELAQIKMYFALA 180  
QY 214 PIATVXHAKSPGTFKLLPDMNKGKKEFLYQTRFLRQVLYLCCQVILQDQCSNIM 273  
Db 181 PIATVXHAKSPGTFKLLPDMNKGKKEFLYQTRFLRQVLYLCCQVILQDQCSNIM 240  
QY 274 LLLGGFTNNMNSRASVYAAHTLAGTSVQNLHWSQAVNSGELRAFQWSETKLNKCN 333  
Db 241 LLLGGFTNNMNSRASVYAAHTLAGTSVQNLHWSQAVNSGELRAFQWSETKLNKCN 300  
QY 334 QPTPVYRVRDMTPTAMWTGGQDWSNPEDVKMLLSEVTNLIYHKNIPPEAHVDFINGL 393  
Db 301 QPTPVYRVRDMTPTAMWTGGQDWSNPEDVKMLLSEVTNLIYHKNIPPEAHVDFINGL 360  
QY 394 DAPHRMYNEIHLMQOQETNLSQGRCEAVL 423  
Db 361 DAPHRMYNEIHLMQOQETNLSQGRCEAVL 390

RESULT 4  
AAU99164  
ID AAU99164 standard; Protein: 398 AA.  
XX AC AAU99164;  
XX 24-SEP-2002 (first entry)  
DE Human lysosomal acid lipase #1.  
XX  
KW Human; enzyme; lysosomal acid lipase; lipid malabsorption illness;  
KW cystic fibrosis; alcoholism; heart disease; heart attack;  
KW Wolman disease; cholesterol ester storage disease; brain injury;  
KW mood disorder; anxiety disorder; thought disorder; volition disorder;  
KW sleep disorder; neurogenic disorder; myopathic disorder; COPD;  
KW obesity; cancer; chronic obstructive pulmonary disease; diabetes;  
KW cardiovascular disorder; Alzheimer's disease; Parkinson's disease;  
KW anorexia; osteoarthritis; central nervous system disorder;  
KW peripheral nervous system disorder.  
OS Homo sapiens.  
XX  
XX WO200236731-A2.  
XX  
XX 10-MAY-2002.  
XX  
XX 30-OCT-2001; 2001WO-EPI2518.  
XX  
XX 31-OCT-2000; 2000US-244170P.  
PR 29-MAY-2001; 2001US-293516P.  
XX  
XX (FARB ) BAYER AG.  
XX  
XX Xiao Y;

```

XX DR WPI; 2002-519248/55.
XX N-PSDB; ABR86569.
XX PT Novel human lysosomal acid lipase polypeptide, useful for treating
XX PT cancer, diabetes, obesity, chronic obstructive pulmonary disease,
XX PT peripheral or central nervous system disorder or cardiovascular
XX disorder.
XX XX
XX PS Claim 25; Fig 2; 126pp; English.
XX CC The invention relates to a purified human lysosomal acid lipase
XX CC polypeptide. Also included are the polynucleotide encoding the
XX CC lipase (or its fragment, derivative, allele or sequence at least 60%
XX CC identical to it), vectors, host cells, a reagent (e.g. an antisense
XX CC oligonucleotide) which binds to the lipase or polynucleotide (used for
XX CC detection and modulating/reducing the lipase activity) and an anti-lipase
XX CC antibody. The lipase and polynucleotide are useful for identifying
XX CC therapeutic agents that either increase or decrease the lipase activity.
XX CC The identified agent, the lipase and polynucleotide are useful for
XX CC treatment of a disease such as lipid malabsorption illness,
XX CC cystic fibrosis, alcoholism, heart disease, heart attack,
XX CC Wolman disease, cholesterol ester storage disease, brain injury,
XX CC mood disorder, anxiety disorder, thought disorder, volition disorder,
XX CC sleep disorder, neurogenic disorder, myopathic disorder, obesity, cancer,
XX CC chronic obstructive pulmonary disease (COPD), diabetes,
XX CC cardiovascular disorder, Alzheimer's disease, Parkinson's disease,
XX CC anorexia, osteoarthritis, a central nervous system disorder and
XX CC a peripheral nervous system disorder. The present sequence is the
XX CC human lysosomal lipase #1.
XX SQ
XX Query Match 88.6%; Score 1991; DB 23; Length 398;
XX Best Local Similarity 100.0%; Pred. No. 9e-194;
XX Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 50 SEIIHQGYPCPEYEYVATEDGYILSVNRIPRGLVQPKKTSRPPVLLQHLGLVGASNWS 109
XX Db SEIIHQGYPCPEYEYVATEDGYILSVNRIPRGLVQPKKTSRPPVLLQHLGLVGASNWS 84
XX
XX QY 110 NLPNNSLGFILADAGFDVWNGNSRGNWSRKHKTLSIDQDEFWAFSYDEMAREDFLPAVIN 169
XX Db NLPNNSLGFILADAGFDVWNGNSRGNWSRKHKTLSIDQDEFWAFSYDEMAREDFLPAVIN 144
XX
XX QY 170 FILQKTQGEKIYVYGYSQGTMTGFIASFSTMPDLAQKIKMYFALAPIATVKHAKSPGKFL 229
XX Db FILQKTQGEKIYVYGYSQGTMTGFIASFSTMPDLAQKIKMYFALAPIATVKHAKSPGKFL 204
XX
XX QY 230 LLPDMIKGLFGKKEFLYQTRFLRQLVYILCGQVILDOICSNIMLLGGFTNNMNSRA 289
XX Db LLPDMIKGLFGKKEFLYQTRFLRQLVYILCGQVILDOICSNIMLLGGFTNNMNSRA 264
XX
XX QY 290 SYVAHTLAGTSVQNTLHWSQAVNSGELRAFDFNGSETKNEKCNQPPVRYRDMTVPT 349
XX Db SYVAHTLAGTSVQNTLHWSQAVNSGELRAFDFNGSETKNEKCNQPPVRYRDMTVPT 324
XX
XX QY 350 AMWTKGODNLSPEDVKMLLSEVTNLIYHKNIPFWAHVDFIWLGDAPHRYNEIHLMOQ 409
XX Db AMWTKGODNLSPEDVKMLLSEVTNLIYHKNIPFWAHVDFIWLGDAPHRYNEIHLMOQ 384
XX
XX QY 410 EETNLSQGRCEAVL 423
XX Db EETNLSQGRCEAVL 398
XX
XX RESULT 5
XX ID AAU99165
XX XX standard; Protein; 371 AA.
XX AC AAU99165;
XX XX
XX DT 24-SEP-2002 (first entry)

```

```

XX DE Human lysosomal acid lipase #2.
XX XX
XX KW Human: enzyme; lysosomal acid lipase; lipid malabsorption illness;
XX KW cystic fibrosis; alcoholism; heart disease; heart attack;
XX KW Wolman disease; cholesterol ester storage disease; brain injury;
XX KW mood disorder; anxiety disorder; thought disorder; volition disorder;
XX KW sleep disorder; neurogenic disorder; myopathic disorder; COPD;
XX KW obesity; cancer; chronic obstructive pulmonary disease; diabetes;
XX KW cardiovascular disorder; Alzheimer's disease; Parkinson's disease;
XX KW anorexia; osteoarthritis; central nervous system disorder;
XX KW peripheral nervous system disorder.
XX OS Homo sapiens.
XX PN WO200236731-A2.
XX XX
XX PD 10-MAY-2002.
XX
XX PF 30-OCT-2001; 2001WO-EP12518.
XX
XX PR 31-OCT-2000; 2000US-244170P.
XX PR 29-MAY-2001; 2001US-293516P.
XX
XX PA (FARB ) BAYER AG.
XX
XX PI Xiao Y;
XX
XX DR WPI; 2002-519243/55.
XX DR N-PSDB; ABR86579.
XX
XX XX
XX PT Novel human lysosomal acid lipase polypeptide, useful for treating
XX PT cancer, diabetes, obesity, chronic obstructive pulmonary disease,
XX PT peripheral or central nervous system disorder or cardiovascular
XX PT disorder.
XX XX
XX PS Claim 25; Fig 5; 126pp; English.
XX
XX CC The invention relates to a purified human lysosomal acid lipase
XX CC polypeptide. Also included are the polynucleotide encoding the
XX CC lipase (or its fragment, derivative, allele or sequence at least 60%
XX CC identical to it), vectors, host cells, a reagent (e.g. an antisense
XX CC oligonucleotide) which binds to the lipase or polynucleotide (used for
XX CC detection and modulating/reducing the lipase activity) and an anti-lipase
XX CC antibody. The lipase and polynucleotide are useful for identifying
XX CC therapeutic agents that either increase or decrease the lipase activity.
XX CC The identified agent, the lipase and polynucleotide are useful for
XX CC treatment of a disease such as lipid malabsorption illness,
XX CC cystic fibrosis, alcoholism, heart disease, heart attack,
XX CC Wolman disease, cholesterol ester storage disease, brain injury,
XX CC mood disorder, anxiety disorder, thought disorder, volition disorder,
XX CC sleep disorder, neurogenic disorder, myopathic disorder, obesity, cancer,
XX CC chronic obstructive pulmonary disease (COPD), diabetes,
XX CC cardiovascular disorder, Alzheimer's disease, Parkinson's disease,
XX CC anorexia, osteoarthritis, a central nervous system disorder and
XX CC a peripheral nervous system disorder. The present sequence is the
XX CC human lysosomal lipase #2.
XX SQ
XX Query Match 68.4%; Score 1536.5; DB 23; Length 371;
XX Best Local Similarity 80.9%; Pred. No. 1.5e-147;
XX Matches 297; Conservative 17; Mismatches 22; Indels 31; Gaps 2;
XX
XX QY 50 SEIIHQGYPCPEYEYVATEDGYILSVNRIPRGLVQPKKTSRPPVLLQHLGLVGASNWS 109
XX Db SEIIHQGYPCPEYEYVATEDGYILSVNRIPRGLVQPKKTSRPPVLLQHLGLVGASNWS 84
XX
XX QY 110 NLPNNSLGFILADAGFDVWNGNSRGNWSRKHKTLSIDQDEFWAFSYDEMAREDFLPAVIN 169
XX Db NLPNNSLGFILADAGFDVWNGNSRGNWSRKHKTLSIDQDEFWAFSYDEMAREDFLPAVIN 144
XX
XX QY 170 FILQKTQGEKIYVYGYSQGTMTGFIASFSTMPDLAQKIKMYFALAPIATVKHAKSPGKFL 229

```





CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic  
CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,  
CC autonomic function disorders such as hypertension and sleep disorders,  
CC neuropsychiatric disorders, psychoactive substance use disorders,  
CC anxiety, and bipolar affective disorder. The present sequence is a  
CC sequence used in a sequence homology alignment with the TANGO/INTERCEPT  
CC sequences of the present invention.

XX  
SQ Sequence 399 AA;

Query Match 57.4%; Score 1289; DB 22; Length 399;  
Best Local Similarity 63.5%; Pred. No. 2.7e-122;  
Matches 235; Conservative 55; Mismatches 80; Indels 0; Gaps 0;

QY 40 AVDPFAMNISEIIIOHGYPCPEEVEVATEDGYILSVNRIPLGLVQPKKTSRPPVLLQHG 99  
DB 28 AVDPETNMVSEIISYWGPFSEELVETEDGYILCLNRIPLHGRKNHSDKGPVVFLLQHG 87  
QY 100 LVGGASWISNLPNNSLGFILADAGFDVWNGSRGNWSRKHKTLSIDQDEFWAFSYDEM 159  
DB 88 LLADSSNWTNLANSSLGFILADAGFDVWNGSRGNWTSRKHKTLSVSQDEFWAFSYDEM 147  
QY 160 ARDPLPAVINFILOKTQGEKIYVYVSGQTTMGTFIAFTMPPELAQIKMYFALAPIATVK 219  
DB 148 AKYDLPASINFILNKTQGEQVYVYVSGQTTIGFIAFSQIPELAKRIKMPFALGPVASVA 207  
QY 220 HAKSPGTFKLLPDMIKLFGKKEFLYOTRFLQVLYLCOVILDOICSNIMLLGGF 279  
DB 208 FCTSPMAKLGRLPDHLKDLFGKEFLPQSAFLKWLGTHTVILKELCGNLCFLCUGF 267  
QY 280 NTNNMNSRASYAAHTLAGTSVQNILHWSQAVNSGELRAFDGMSKNEKCNQPTPVR 339  
DB 268 NERNLNRSVDVYTHSPAGTSVQNMLHWSQAVKFKQAFDNGSSAKNYFHYNQSYPT 327  
QY 340 YRVDMVTPTAMTGGQDMLSNPDEVKMLSEVNTLIYHNIPENAHVDFIWLGDAPHRM 399  
DB 328 YNVKMDLVPTAVMSGGHDVADYDYNILLTQITNLVFHESIPWEHLDFIWLGDAPWRL 387  
QY 400 YNEIHLMOQ 409  
DB 388 YNKIILMRK 397

#### RESULT 8

AAO18227  
ID AAO18227 standard; Protein; 392 AA.

XX  
AC AAO18227;

18-SEP-2002 (first entry)

Human lysosomal acid lipase related protein.

XX Human; lysosomal acid lipase; chromosome 10; lipase; Wolman disease;  
KW cholesteryl ester storage disease.

XX Homo sapiens.

XX US6387680-B1.

XX 14-MAY-2002.

XX 29-MAR-2001; 2001US-0820001.

XX 29-MAR-2001; 2001US-0820001.

XX (PEKE ) PE CORP NY.

XX Merkulov GV, Ketchum KA, Di Francesco V, Beasley EM;

XX WPI; 2002-478445/51.

XX Isolated nucleic acids encoding a lysosomal acid lipase protein useful

PT for the prevention, diagnosis and treatment of severe infantile-onset  
PT Wolman disease and late-onset cholesteryl ester storage disease -  
XX  
XX Disclosure; Column 57-60; 68pp; English.

CC The present invention provides the protein, coding and genomic sequences  
CC of a human lysosomal acid lipase. The sequences can be used in the  
CC identification of modulators of lipase activity in cells and tissues that  
CC express the lipase, particularly the severe infantile-onset Wolman  
CC disease and the milder late-onset cholesteryl ester storage disease  
CC (CESD), which are caused by mutations in different parts of the lysosomal  
CC acid lipase (LIPA) gene. The present sequence is a protein shown in the  
CC exemplification of the invention.

XX  
SQ Sequence 392 AA;

Query Match 57.1%; Score 1282; DB 23; Length 392;  
Best Local Similarity 63.8%; Pred. No. 1.4e-121;  
Matches 234; Conservative 53; Mismatches 80; Indels 0; Gaps 0;

QY 40 AVDPFAMNISEIIIOHGYPCPEEVEVATEDGYILSVNRIPLGLVQPKKTSRPPVLLQHG 99  
DB 26 AVDPETNMVSEIISYWGPFSEELVETEDGYILCLNRIPLHGRKNHSDKGPVVFLLQHG 85  
QY 100 LVGGASWISNLPNNSLGFILADAGFDVWNGSRGNWSRKHKTLSIDQDEFWAFSYDEM 159  
DB 86 LLADSSNWTNLANSSLGFILADAGFDVWNGSRGNWTSRKHKTLSVSQDEFWAFSYDEM 145  
QY 160 ARDPLPAVINFILOKTQGEKIYVYVSGQTTMGTFIAFTMPPELAQIKMYFALAPIATVK 219  
DB 146 AKYDLPASINFILNKTQGEQVYVYVSGQTTIGFIAFSQIPELAKRIKMPFALGPVASVA 205  
QY 220 HAKSPGTFKLLPDMIKLFGKKEFLYOTRFLQVLYLCOVILDOICSNIMLLGGF 279  
DB 206 FCTSPMAKLGRLPDHLKDLFGKEFLPQSAFLKWLGTHTVILKELCGNLCFLCUGF 265  
QY 280 NTNNMNSRASYAAHTLAGTSVQNILHWSQAVNSGELRAFDGMSKNEKCNQPTPVR 339  
DB 266 NERNLNRSVDVYTHSPAGTSVQNMLHWSQAVKFKQAFDNGSSAKNYFHYNQSYPT 325  
QY 340 YRVDMVTPTAMTGGQDMLSNPDEVKMLSEVNTLIYHNIPENAHVDFIWLGDAPHRM 399  
DB 326 YNVKMDLVPTAVMSGGHDVADYDYNILLTQITNLVFHESIPWEHLDFIWLGDAPWRL 385  
QY 400 YNEIHL 405  
DB 386 YNKIIL 392

#### RESULT 9

AAU30498

ID AAU30498 standard; Protein; 449 AA.

XX  
AC AAU30498;

18-DEC-2001 (first entry)

Novel human secreted protein #989.

XX Human; vaccination; gene therapy; nutritional supplement;

XX stem cell proliferation; haematopoiesis; nerve tissue regeneration;

XX immune suppression; immune stimulation; anti-inflammatory; leukaemia.

XX Homo sapiens.

XX WO200179449-A2.

XX 25-OCT-2001.

XX 16-APR-2001; 2001WO-US08656.

XX 18-APR-2000; 2000US-0552929.

XX 26-JAN-2001; 2001US-0770160.

XX PA (HYSE-) HYSEQ INC.  
XX PI Tang YT, Liu C, Drmanac RT;  
XX DR WPI: 2001-611725/70.  
XX XX  
XX PT Nucleic acids encoding a range of human polypeptides, useful in genetic  
XX PT vaccination, testing and therapy -  
XX PS Claim 20; Page 301; 765pp; English.  
XX XX  
XX CC The invention relates to novel human secreted polypeptides. The  
XX CC polypeptides and antibodies to the polypeptides are useful for  
XX CC determining the presence of or predisposition to a disease associated  
XX CC with altered levels of polypeptide. The polypeptides are also useful for  
XX CC identifying agents (agonists and antagonists) that bind to them. Cells  
XX CC expressing the proteins are useful for identifying a therapeutic agent  
XX CC for use in treatment of a pathology related to aberrant expression or  
XX CC physiological interactions of the polypeptide. Vectors comprising  
XX CC the nucleic acids encoding the polypeptides and cells genetically  
XX CC engineered to express them are also useful for producing the proteins.  
XX CC The proteins are useful in genetic vaccination, testing and  
XX CC therapy, and can be used as nutritional supplements. They may be used to  
XX CC increase stem cell proliferation; to regulate haematopoiesis; and in  
XX CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;  
XX CC immune suppression and/or stimulation; as anti-inflammatory agents; and  
XX CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid  
XX CC sequences of novel human secreted proteins of the invention.

XX SQ Sequence 449 AA;

Query Match 53.0%; Score 1191.5; DB 22; Length 449;  
Best Local Similarity 57.4%; Pred. No. 2.7e-112;  
Matches 230; Conservative 59; Mismatches 109; Indels 3; Gaps 2;  
Qy 12 SHRMEMULLILVAYFQFNVSVMPTK--AVDPEAFNISEIIOHQGYPCEEVEVATED 69  
Dy 47 SSRKMRVGLVCLVLTSLHSESGGKLTAVDPETNNVSEIISYMGFPSEEVIVTED 106  
Qy 70 GYILSVNRIPLGLVQPKTKSRPVVLLQHLGVGGASNWSLNPLNSLGLFADAGFDVYM 129  
Dy 107 GTILCLNRIPIGRKNHSDKPKVPVFLQHLGLADSSNVTNLGNSRLGLFADADAIDYM 166  
Qy 130 GNSRGNASRRKHTLSIDQDEFWAFSYDEMARFDLPVAVINFLAQTKGQEKIYVYGSQGT 189  
Dy 167 GNTRGNTWSPKHTLSVSQDEFWAFSYDEMAKYDLPASINFLNKTGQEQYVYVGHSGGT 226  
Qy 190 TWGFTAFSTMPDLAQKIMYFALAPIATVVKHAKSPGTFFLLPDMIMKGLFGKKEFLYQT 249  
Dy 227 TIGTFAFQMLELAKGLKMFANGPVASVAFCTSPMAKLGRLPDHLKDLFGDEEFLPOS 286  
Qy 250 RFLROL-VIYLCGGVILDQICSNIMLLGGENTNNMNSRASVYAAHTLAGTSVQNILHW 308  
Dy 287 AFWKVAGVPHLATHYVILKELCNLCFLICGFGNERLNLSRDVYVTHSPAGTFVQNMXXH 346  
Qy 309 SOAVNSGELRAFDMGSETKNEKCNQPTPVRYRVRDMVTPTAMTGGODWLSNPDVKNL 368  
Dy 347 SOAVKFORQAFDQSSAKNFYHYNQSYPTYNVYKMDLYPXTAVTGGHDLVDYGVNIX 406  
Qy 369 LSEVTNLIYHKNIPWAHVDFIWLGDAPHRYMNEIHLMOQ 409  
Dy 407 LTOITNLVHESIPWEHLDPIWGLDAPRWLYNKIINLRK 447

RESULT 10

AAAR37302  
ID AAR37302 standard; Protein: 398 AA.

XX AC AAR37302;

XX XX  
XX DT 20-SEP-1993 (first entry)

DE RGL precursor.  
XX XX Rabbit gastric lipase; RGL; PRGLN2.1; fat; bioconversion;  
KW hydrolysis; transesterification.  
XX OS Oryctolagus cuniculus.  
XX XX  
XX FH Key Location/Qualifiers  
FT Protein 23..398  
FT /note= "claim 1; page 10-11"  
XX XX  
XX PN EP542629-A.  
XX XX 19-MAY-1993.  
XX PD 12-NOV-1992; 92EP-0403055.  
XX PF 13-NOV-1991; 91FR-0013948.  
XX PR (LJOJ ) INST RECH JOUVEINAL.  
XX PA Benicourt C, Blanchard C, Junien J;  
XX PI WPI: 1993-161080/20.  
XX DR N-PSDB: AAQ42310.  
XX XX  
XX PT Rabbit gastric lipase, its precursor and their DNA - useful for  
XX PT treating conditions linked to gastric lipase deficiency, such as  
XX PT mucoviscidiosis and pancreatic exocrine insufficiency  
XX PS Claim 1; Fig 7; 31pp; French.  
XX SQ Sequence 398 AA;

Query Match 51.9%; Score 1166; DB 14; Length 398;  
Best Local Similarity 54.6%; Pred. No. 9e-110;  
Matches 219; Conservative 70; Mismatches 102; Indels 10; Gaps 3;  
Qy 17 MWLLILVAYMFQFNVSVMPTKAVDPEAFNISEIIOHQGYPCEEVEVATEDG 70  
Dy 1 MWVLFVAAALISA-LGTHGLFGKSAPT---NPEVNMNLSQMSIYWGYPSEYEVATEDG 56  
Qy 71 YILSVNRIPLGLVQPKTKSRPVVLLQHLGVGGASNWSLNPLNSLGLFADAGFDVYM 130  
Dy 57 YILEVNRIPYKKNNGRGPVFLQHLGLASASNWSLNPLNSLGLFADAGVWLG 116  
Qy 131 NSRGNASRRKHTLSIDQDEFWAFSYDEMARFDLPVAVINFLAQTKGQEKIYVYGSQGT 190  
Dy 117 NSRGNASRRKHTLSIDQDEFWAFSYDEMARFDLPVAVINFLAQTKGQEKIYVYGSQGT 176  
Qy 191 MGFIAFSTMPDLAQKIMYFALAPIATVVKHAKSPGTFFLLPDMIMKGLFGKKEFLYQT 250  
Dy 177 IGFIAFSTMPDLAQKIMYFALAPIATVVKHAKSPGTFFLLPDMIMKGLFGKKEFLYQT 236  
Qy 251 FLRQLVIYLCGGVILDQICSNIMLLGGENTNNMNSRASVYAAHTLAGTSVQNILHW 310  
Dy 237 FDQFLATQVCSRETNLVICSNALFICGFDNSANLSRLDYYVSHNPACTSVQNMLHW 296  
Qy 311 AVHSGELRAFDMGSETKNEKCNQPTPVRYRVRDMVTPTAMTGGODWLSNPDVKNL 370  
Dy 297 AVKSGNFOAFNGWSPQNVHFNQPTPVRYRVRDMVTPTAMTGGODWLSNPDVKNL 356  
Qy 371 EYVNLIIYHKNIPWAHVDFIWLGDAPHRYMNEIHLMOQ 411  
Dy 357 KLSNLIYHKEILLPYNHLDFIWMNAPQEVYNEIISMAKDK 397

RESULT 11  
AAB66068  
ID AAB66068 standard; Protein; 221 AA.  
XX  
AC AAB66068;  
XX  
DT 30-MAR-2001 (first entry)  
XX  
DE Human TANGO 294 extracellular domain.  
XX  
KW TANGO protein; INTERCEPT protein; neurological disorder;  
KW central nervous system; focal brain disorder; bipolar affective disorder;  
KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;  
KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;  
KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;  
KW neuropsychiatric; psychoactive substance use; anxiety.  
XX  
OS Homo sapiens.  
XX  
PN WO200077239-A2.  
XX  
PD 21-DEC-2000.  
XX  
PF 24-MAY-2000; 2000WO-US14858.  
XX  
PR 14-JUN-1999; 99US-0333159.  
XX  
PA (MILL-) MILLENNIUM PHARM INC.  
XX  
PI McCarthy SA, Fraser CC, Sharp JD, Barnes TM;  
XX  
DR WPI; 2001-032313/04.  
XX  
PT TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for  
PT screening assays, and diagnostic assays and for the treatment of  
PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's  
PT disease.  
XX  
PS Claim 8; Page 326; 359pp; English.  
XX  
CC The present invention relates to TANGO or INTERCEPT proteins and coding  
CC sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,  
CC AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding  
CC sequences are useful for the treatment of neurological disorders such as  
CC central nervous system (CNS) disorders, CNS-related disorders, focal  
CC brain disorders, global-diffuse cerebral disorders and other  
CC neurological and cerebrovascular disorders. The CNS disorders include  
CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic  
CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,  
CC autonomic function disorders such as hypertension and sleep disorders,  
CC neuropsychiatric disorders, psychoactive substance use disorders,  
CC anxiety, and bipolar affective disorder.  
XX  
SQ Sequence 221 AA;  
Query Match 51.8%; Score 1165; DB 22; Length 221;  
Best Local Similarity 100.0%; Pred. No. 4.6e-110;  
Matches 221; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 34 VHMPTKAVDPAEFMNIIEIIHQGYPCPEEYEVATEDGYILSVNRPGLVQPKTGRPV 93  
DB 1 VHMPTKAVDPAEFMNIIEIIHQGYPCPEEYEVATEDGYILSVNRPGLVQPKTGRPV 60  
QY 94 VLLQHLVGGASNWSNLPNNSLGLFIADAGFVWNGSRGNWSRKHKTLSDQDEFWA 153  
DB 61 VLLQHLVGGASNWSNLPNNSLGLFIADAGFVWNGSRGNWSRKHKTLSDQDEFWA 120  
QY 154 FSYDEMARFDPVAVINFILOKTGOEKIYYGYSGQTTMGFIAPSTMPPELAOKIKMYFALA 213  
DB 121 FSYDEMARFDPVAVINFILOKTGOEKIYYGYSGQTTMGFIAPSTMPPELAOKIKMYFALA 180  
QY 214 PIATVKHAKSPGTFKLLLPDMMIKGLFGKKEFLYQTRFLRQ 254

DB 181 PIATVKHAKSPGTFKLLLPDMMIKGLFGKKEFLYQTRFLRQ 221  
RESULT 12  
AAP50322  
ID AAP50322 standard; protein; 395 AA.  
XX  
AC AAP50322;  
XX  
DT 17-JAN-1992 (first entry)  
XX  
DE Rat lingual lipase protein.  
XX  
KW Lingual lipase; enzyme; EC-3.1.1.3; ss.  
XX  
OS Rattus rattus.  
XX  
PN GB2142337-A.  
XX  
PD 16-JAN-1985.  
XX  
PF 29-JUN-1984; 84GB-0016581.  
XX  
PR 01-JUL-1983; 83GB-0017989.  
XX  
PR 05-SEP-1983; 83GB-0023759.  
XX  
PA (CELL-) CELLTECH LTD.  
XX  
PI Carey NH, Williamson R;  
XX  
DR WPI; 1985-014450/03.  
XX  
DR N-PSDB; AAN50385.  
XX  
PT New lingual lipase protein for treatment of lipase deficiency - also  
PT new pre-lingual lipase protein and related products  
XX  
PS Disclosure; Fig 5; 15pp; English.  
XX  
CC This protein may be expressed in a transformant host organism  
CC and may be used for the treatment of lipase deficiency.  
XX  
SQ Sequence 395 AA;  
Query Match 51.5%; Score 1156.5; DB 6; Length 395;  
Best Local Similarity 54.7%; Pred. No. 8.2e-109;  
Matches 216; Conservative 69; Mismatches 109; Indels 1; Gaps 1;  
QY 17 MWLLILVAYMFQ-RNVNSVHMPTKAVDPAEFMNIIEIIHQGYPCPEEYEVATEDGYILSV 75  
DB 1 MWLLILVAYMFQ-RNVNSVHMPTKAVDPAEFMNIIEIIHQGYPCPEEYEVATEDGYILSV 60  
QY 76 NRIPRGLVQPKTGRPVVLLQHLVGGASNWSNLPNNSLGLFIADAGFVWNGSRGN 135  
DB 61 YRIPHGKNNSENIGKRPVVYLQHLGLIASATNWIANLPNNSLAFMLADAGYDVLGNSRGN 120  
QY 136 AWSRKHKTLSDQDEFWAFSYDENAREFDPVAVINFILOKTGOEKIYYGYSGQTTMGFIA 195  
DB 121 TWSRKNYYSPDSVEFWAFSDEMAKIDLPATINFIYQKTGOEKIHYVGHQSQRITGFA 180  
QY 196 FSTMPPELAOKIKMYFALAPATVTKHAKSPGTFKLLLPDMMIKGLFGKKEFLYQTRFLRQ 255  
DB 181 FSTNPTLAKKIKTYALAPVATKYTOSPLKKISFIPTFLFKLMFGKMKFLPHYFDDFL 240  
QY 256 VYLQCGVILQDQICSNIMLLGGFTNNMNSRASVVAHTLAGTSVQNIHLHWSQAVNSG 315  
DB 241 GTEVCSREVLJLLCSNTLFIYQKGFDPKKNLNVSRFDVYLGHNPAGTSVQDFLWLAQLVRSG 300  
QY 316 ELRAFDPGSETKNLEKCNQPTPVRYRDMTPVPTAMWTGGODWLSNPDVAKMLSEVTNL 375  
DB 301 KFOAFNMGSPSQSNMLHYNQKTPPEYDVSAMTPVPVAVWNGNDILADPODVAMLLPKLSNL 360  
QY 376 IYHKNIPWAHVDFIWLGLDAPHRMVNEIHLMQOE 410



PT preventing immune system disorders (e.g. Crohn's disease), neurological disorders (e.g. Parkinson's disease), or cell proliferative disorders (e.g. cancers) -

PS Claim 48; Page 111-112; 122pp; English.

XX The present invention relates to the isolation of human lipid metabolism enzymes (LMM) designated LMM-1 to LMM-6, and the polynucleotide sequences encoding them. The LMM polypeptides, polynucleotides, agonists and antagonists are useful for diagnosing, treating or preventing disorders associated with aberrant expression of LMM, particularly immune system disorders (e.g. acquired immunodeficiency syndrome (AIDS), thymic hypoplasia, anaemia, asthma or Crohn's disease), neurological disorders (e.g. epilepsy, Huntington's disease, dementia or Parkinson's disease), developmental disorders (e.g. Down's syndrome) or cell proliferative disorders (e.g. cancers including adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma or sarcoma). The present sequence represents human LMM-4.

XX Sequence 399 AA;

Query Match 50.9%; Score 1143; DB 23; Length 399;

Best Local Similarity 54.2%; Pred. No. 2e-107;

Matches 213; Conservative 71; Mismatches 101; Indels 8; Gaps 2;

QY 18 WLLILVA-YMFORNVNSVHMTKAVDPAPMNTSEIIHQGYPCPEEYEVATEDGYILSVN 76

DB 10 WMLLGSMYGDKGNN-----ANPEANNISQIISYWGYPYEVYDVTTKDGYILGIY 62

QY 77 RIRPGLVOPKTSRPPVLLQHLGVGASNWSNLPNNSLGFILADGFDVWNGNSRGNA 136

DB 63 RIRPGRGCPGRTAPKPAVYLQHLGLIASANWICNLPNNSLAFLLADSGYDVLGNSRGNT 122

QY 137 WSRKHKLTDQDEFWAFSDENARFDLPVINFILOKTGOEKIYVYSGTGMGTAF 196

DB 123 WSRKHKLSPSEIYWFSLDEMAKYDLPATINFIEIKTGKRLYVYVGHSGTGMGTAF 182

QY 197 STNPELAKIKMYFALAPIATVVKHAKSPGTFKLLPDMIMKGLFGKKEFLYQTRFLRLV 256

DB 183 STNPELAKIKIFALAPVTVVYKTSQPMKLLTSLRRVVKVLFSGDKMFHPTLFDQFIA 242

QY 257 IYLCGGVILQICSNIMLLGGFTNNMNSRASVYAAHTLAGTSVQNLHWSQAVNSGE 316

DB 243 TKVCRNKLFRICSNFLFTLSGFDPPQNLNMRSLDVLVSHNPAGTSVQNLHWSQAVNSGQ 302

QY 317 LRAFDMGSETKNLEKCHQPTVRYVRVDMVPTAMWTGGODWLNPNEDVKMLLSEVTNLI 376

DB 303 LQAFDWNQSDONMMHFQLTPLPLNITKMEVPTAINWGGQDIVADPKDVENLLPQIANLI 362

QY 377 YHKNIPWAHVDFIWLGDAPHRYMNEIHLMOQ 409

DB 363 YKLLPHYNHVDFLGEDAPOEIVODLILME 395

RESULT 15

AAAR56870

XX AAR56870 standard; Protein; 379 AA.

AC AAR56870;

XX 20-FEB-1995 (first entry)

DT Canine gastric lipase.

DE Canine gastric lipase; CGI; dog; fat; mucoviscidosis; enzyme;

XX Canine gastric lipase; CGI; dog; fat; mucoviscidosis; enzyme;

KW blooconversion; exocrine pancreatic insufficiency.

XX Canis familiaris.

OS WO9413816-A.

XX 23-JUN-1994.

XX

PF 16-DEC-1993; 93WO-FR01260.

XX 16-DEC-1992; 92PR-0015201.

PR (LJOU ) INST RECH JOUVEINAL.

PA Benicourt C, Blanchard C, Junien J;

XX WPI: 1994-21789C/26.

XX N-PSDB; AAQ68386.

XX Recombinant canine gastric lipase and nucleic acid encoding it -

PT are used for improving absorption of ingested fat, treating

XX mucoviscidosis etc. and in enzymatic bio-conversions

XX Claim 13; Fig 9A; 52pp; French.

XX The sequence given below is the sequence of figure 9A, altered

CC according to the amendments described on page 2 of the appended

CC letter.

CC CGI is used to improve absorption of ingested fat, in healthy and

CC sick patients (e.g. having altered levels of gastric lipase); to

CC treat conditions associated with insufficiency (or lack) of lipases,

CC esp. mucoviscidosis or exocrine pancreatic insufficiency and partic.

CC where immobilised, for bioconversions, e.g. hydrolysis or

CC transesterification (other mammalian gastric lipases, or derivs.,

CC can be used in this application).

XX Sequence 379 AA;

QY Query Match 50.8%; Score 1141.5; DB 15; Length 379;

DB Best Local Similarity 56.5%; Pred. No. 2.6e-107;

Matches 210; Conservative 67; Mismatches 90; Indels 5; Gaps 3;

QY 37 PTKAVDEAFANISEIITQHQGYPCPEEYEVATEDGYILSVNIPRGLVQPKTGRPPVLL 96

DB 7 PT---NPEVTNISMITYMGYPAEYEVATEDGYILGIDRIPYGRKNSENIGRRPVAFL 63

QY 97 QHGLVGGASNWSNLPNNSLGFILADGFDVWNGNSRGNAWSRKHKLTDQDEFWAFS 156

DB 64 QHGLASATNWSNLPNNSLAFILADGFDVWNGNSRGNTWARRNLYSPDSVEFNAFS 123

QY 157 DEMARFDLPVINFILOKTGOEKIYVYSGTGMGTAFSTMPPELAKIKMYFALAPIA 216

DB 124 DEMAKYDLPATIDFLAKTGQDKLHYVGHSGTGTIGTAFSTNPKLAKRIKTFYALAPVA 183

QY 217 TVKHAKSPGTFKLLPDMIMKGLFGKKEFLYQTRFLRO-LVIYLCGGVILQICSNIMLL 275

DB 184 TVKTEILLNKLMLVPSFLFKLIFGNKIF-YPHHFFDQFLATEVCSRETVDLLCSNALFI 242

QY 276 LGGFNTNNMNSRASVYAAHTLAGTSVQNLHWSQAVNSGELRAFDMGSETKNLEKNQ 335

DB 243 ICGEDTNLNSRLDVLVSHNPAGTSVQNLHWSQAVNSGELRAFDMGSETKNLEKNQ 302

QY 336 TPVRYVRDMVPTAMWTGGODWLNPNEDVKMLLSEVTNLIYHKNIPWAHVDFIWLGD 395

DB 303 MPYVNTLDMHVPITAVWNGNDLADPHDVLKLLSKPLNLIYHKKIPPYNHLDFIWMADA 362

QY 396 PHRYMNEIHLMO 407

DB 363 POAVYNEIVSMW 374

Search completed: February 3, 2003, 13:48:54

Job time : 41 secs

GenCore version 5.1.3  
Copyright (C) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:50:05 ; Search time 13 Seconds  
(without alignments)  
656.579 Million cell updates/sec

Title: US-09-333-159-47

Perfect score: 2247

Sequence: 1 MLETLSRQWVSHRMWMLL.....IHLMOQETNLSQGRCEAVL 423

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 122226 seqs, 20178551 residues

Total number of hits satisfying chosen parameters: 122226

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

Published Applications\_AA.\*

- 1: /cgn2\_5/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 2: /cgn2\_5/ptodata/2/pubpaa/PTCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_5/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_5/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_5/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_5/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_5/ptodata/2/pubpaa/PTCTUS\_PUBCOMB.pep.\*
- 8: /cgn2\_5/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_5/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 10: /cgn2\_5/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_5/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 12: /cgn2\_5/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_5/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 14: /cgn2\_5/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2247	100.0	423	9 US-10-042-431-47	Sequence 47, Appl
2	2174	96.8	409	10 US-09-835-996A-21	Sequence 21, Appl
3	2076	92.4	390	9 US-10-042-431-49	Sequence 49, Appl
4	1991	88.6	374	10 US-09-811-825-4	Sequence 4, Appl
5	1991	88.6	398	10 US-09-811-825-2	Sequence 2, Appl
6	1289	57.4	399	9 US-10-042-431-41	Sequence 41, Appl
7	1282	57.1	392	12 US-10-003-302-4	Sequence 4, Appl
8	1165	51.8	221	9 US-10-042-431-50	Sequence 50, Appl
9	1119	49.8	398	9 US-10-042-431-75	Sequence 75, Appl
10	1090	48.5	398	9 US-10-056-744B-2	Sequence 2, Appl
11	1031	45.9	395	12 US-10-003-302-2	Sequence 2, Appl
12	783	34.8	144	9 US-10-042-431-52	Sequence 52, Appl
13	400	17.8	127	9 US-10-056-744B-5	Sequence 5, Appl
14	171	7.6	33	9 US-10-042-431-48	Sequence 48, Appl
15	128	5.7	25	9 US-10-042-431-51	Sequence 51, Appl
16	109.5	4.9	305	10 US-09-815-242-12337	Sequence 12337, A
17	103	4.6	222	10 US-09-815-242-5276	Sequence 5276, Ap
18	102	4.5	346	9 US-10-027-805-34	Sequence 34, Appl
19	102	4.5	346	10 US-09-903-410-34	Sequence 34, Appl

20	39	4.4	678	9 US-09-738-626-6025	Sequence 6025, Ap
21	96.5	4.3	287	12 US-10-080-644-4	Sequence 4, Appl
22	96.5	4.3	1275	9 US-10-025-201-3	Sequence 3, Appl
23	95.5	4.3	917	9 US-10-001-835-189	Sequence 189, App
24	33	4.1	637	9 US-09-738-626-4721	Sequence 4721, Ap
25	32	4.1	748	10 US-09-815-242-10278	Sequence 10278, A
26	32	4.1	748	10 US-09-815-242-13880	Sequence 13880, A
27	31	4.0	298	10 US-09-950-368-6	Sequence 6, Appl
28	39	4.0	342	10 US-09-896-578-2	Sequence 2, Appl
29	87.5	3.9	297	12 US-10-080-644-7	Sequence 7, Appl
30	86.5	3.8	624	9 US-09-738-626-4289	Sequence 4289, Ap
31	86.5	3.8	867	9 US-09-839-894-6	Sequence 6, Appl
32	35	3.8	843	10 US-09-873-737A-2	Sequence 2, Appl
33	84.5	3.8	338	9 US-10-027-805-40	Sequence 40, Appl
34	84.5	3.8	338	10 US-09-903-410-40	Sequence 40, Appl
35	34	3.7	389	10 US-09-967-645-1	Sequence 1, Appl
36	82.5	3.7	331	9 US-09-738-626-3832	Sequence 3832, Ap
37	82.5	3.7	485	9 US-09-918-543-2	Sequence 2, Appl
38	82.5	3.7	485	9 US-09-795-211-1	Sequence 1, Appl
39	82.5	3.7	485	10 US-09-769-864-1	Sequence 1, Appl
40	82.5	3.7	485	10 US-09-769-864-7	Sequence 7, Appl
41	82.5	3.7	485	10 US-09-854-346-2	Sequence 2, Appl
42	82.5	3.7	485	10 US-09-902-188A-1	Sequence 1, Appl
43	82.	3.6	330	9 US-10-051-643-30	Sequence 30, Appl
44	82	3.6	330	9 US-09-880-505-30	Sequence 30, Appl
45	81.5	3.6	197	9 US-09-918-543-13	Sequence 13, Appl

#### ALIGNMENTS

##### RESULT 1

US-10-042-431-47  
; Sequence 47, Application US/10042431  
; Publication No. US20020182675A1

GENERAL INFORMATION:  
; APPLICANT: MCCARTHY, Sean A

; APPLICANT: BARNES, Thomas M  
; APPLICANT: FRASER, Christopher C

; APPLICANT: SHARP, John D  
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,

; PREVENTIVE, THERAPEUTIC, AND OTHER USES  
; FILE REFERENCE: 10147-602  
; CURRENT APPLICATION NUMBER: US/10042.431

; CURRENT FILING DATE: 2001-10-25  
; PRIOR FILING DATE: 2001-10-25

; PRIOR FILING DATE: 1999-06-14  
; PRIOR APPLICATION NUMBER: US 09/333,159

; PRIOR FILING DATE: 1999-06-14  
; PRIOR APPLICATION NUMBER: US 09/578,063

; NUMBER OF SEQ ID NOS: 79  
; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 47  
; LENGTH: 423

; TYPE: PRT  
; ORGANISM: Homo sapiens

US-10-042-431-47

Query Match 100.0%; Score 2247; DB 9; Length 423;

Best Local Similarity 100.0%; Pred. No. 4e-204;

Mismatches 0; Indels 0; Gaps 0;

QY 1 MLETLSRQWVSHRMWMLLILVAYMFQNVNSVHMPTKAVDPFAFNISIIHQGYPC 60

Db 1 MLETLSRQWVSHRMWMLLILVAYMFQNVNSVHMPTKAVDPFAFNISIIHQGYPC 60

QY 61 EEEVATEDGYILSVNRIPGLVQPKTGRPVVLLQHLGVLGGASNWSLNPLNSLGLFIL 120

Db 61 EEEVATEDGYILSVNRIPGLVQPKTGRPVVLLQHLGVLGGASNWSLNPLNSLGLFIL 120

QY 121 ADAGFDVWMCNSRGNWSRKHKTLSIDQDEFWAFSYDEMAREFDLPVAFINFLQKTGQEKI 180

Db 121 ADAGFDVWMCNSRGNWSRKHKTLSIDQDEFWAFSYDEMAREFDLPVAFINFLQKTGQEKI 180

Qy	181	YYVYSGQTGMGFAESTMPELAQKTKMYTALAPIATVKHAKSPGTKFLLLPDMIMIKGLF	240
Db	181	YYVYSGQTGMGFAESTMPELAQKTKMYTALAPIATVKHAKSPGTKFLLLPDMIMIKGLF	240
Qy	241	GKKEFLYQTRFLRQLVYLICGVILDOICSNTMLLGGFTNNMNKSRSVAAHTLAGT	300
Db	241	GKKEFLYQTRFLRQLVYLICGVILDOICSNTMLLGGFTNNMNKSRSVAAHTLAGT	300
Qy	301	SVQNIIHWQAQVNSGELRAFDFMGSETKNLEKCNPQTPPVRYVRDMTVP TAMTNGGDWLS	360
Db	301	SVQNIIHWQAQVNSGELRAFDFMGSETKNLEKCNPQTPPVRYVRDMTVP TAMTNGGDWLS	360
Qy	361	NPEDVMKLSEVTNLIYHKNIPEWAHVDFIWLGLDAPHRMYNETIIHLMQEETNLSSGRCE	420
Db	361	NPEDVMKLSEVTNLIYHKNIPEWAHVDFIWLGLDAPHRMYNETIIHLMQEETNLSSGRCE	420
Qy	421	AVL 423	
Db	421	AVL 423	

## RESULT 2

```

US-09-835-996A-21
: Sequence 21, Application US/09835996A
: Patent No. US20020142953A1
: GENERAL INFORMATION:
: APPLICANT: Ballinger, Dennis
: APPLICANT: Loeb, Debra
: APPLICANT: Montgomery, Julie
: APPLICANT: Tang, Y. Tom
: APPLICANT: Zhou, Ping
: APPLICANT: Goodrich, Ryle
: APPLICANT: Liu, Chenghua
: APPLICANT: Asundi, Vinod
: APPLICANT: Zhao, Qing
: APPLICANT: Wehrman, Tom
: APPLICANT: Drmanac, Radoje
: APPLICANT: Ren, Feiyang
: APPLICANT: Qian, Xiaohong
: APPLICANT: Wang, Duanrui
: TITLE OF INVENTION: MATERIALS AND METHODS RELATING TO LIPID METABOLISM
: FILE REFERENCE: 28110/35915A
: CURRENT APPLICATION NUMBER: US/09/835,996A
: CURRENT FILING DATE: 2001-04-16
: PRIOR APPLICATION NUMBER: US 60/197,137
: PRIOR FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: US 09/714,936
: PRIOR FILING DATE: 2000-11-17
: PRIOR APPLICATION NUMBER: US 09/667,298
: PRIOR FILING DATE: 2000-09-22
: PRIOR APPLICATION NUMBER: US 09/631,451
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: US 09/598,042
: PRIOR FILING DATE: 2000-06-20
: NUMBER OF SEQ ID NOS: 45
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 21
: LENGTH: 409
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-835-996A-21

```

## RESULTS

```

US-10-042-431-49
; Sequence 49, Application US/10042431
; Publication No. US20020182675A1
; GENERAL INFORMATION:
; APPLICANT: MCCARTHY, Sean A
; APPLICANT: BARNES, Thomas M
; APPLICANT: FRASER, Christopher C
; APPLICANT: SHARP, John D
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
; THERAPEUTIC, AND OTHER USES
; FILE REFERENCE: 10147-602
; CURRENT APPLICATION NUMBER: US/10/042,431
; CURRENT FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 49
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-042-431-49

```

Query Match	92.4%;	Score 2076;	DB 9;	Length 390;
Best Local Similarity	100.0%;	Pred. No. 4.8e-188;		
Matches 390;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	34	VHMPKAVDPEAFNNISIIQHOGYPCPEEYEVATEDGYILSVNRI	PGLVOPKKTGSRPV	93
Db				
	1	VHMPKAVDPEAFNNISIIQHOGYPCPEEYEVATEDGYILSVNRI	PGLVOPKKTGSRPV	60
Qy	94	VLLQHGVLGGASWNISLNPNSLGLFIADAGDFVWNGSRGNAN	SRKHKTULSI	153
Db				
	61	VLLQHGVLGGASWNISLNPNSLGLFIADAGDFVWNGSRGNAN	SRKHKTULSI	120
Qy	154	FSYDEMARFDLPVINFILQKTGGEKIIYVYGVSOGTGMGFAT	STMPPELAQIK	213
Db				
	121	FSYDEMARFDLPVINFILQKTGGEKIIYVYGVSOGTGMGFAT	STMPPELAQIK	180
Qy	214	PIATVKHAKSPGTFFLLPDMNMIKGLFGKKEFLYQTRFL	RQLRVIIYLCQV	273
Db				
	181	PIATVKHAKSPGTFFLLPDMNMIKGLFGKKEFLYQTRFL	RQLRVIIYLCQV	240
Qy	274	LLLGGFNTNNMNSRASYAAHTLAGTSVQNILHWSQAVNS	GEURAFDWSG	333
Db				
	241	LLLGGFNTNNMNSRASYAAHTLAGTSVQNILHWSQAVNS	GEURAFDWSG	300
Qy	334	QPTPVRYRVFDMTVPYTAWMTGGQDWLSPNPDVKMLLS	EVNTNLIYHKNI	393

Query Match	92.4%;	Score 2076;	DB 9;	Length 390;
Best Local Similarity	100.0%;	Pred. No. 4.8e-188;		
Matches 390;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	34	VHMPKAVDPEAFNNISIIQHOGYPCPEEYEVATEDGYILSVNRI	PGLVOPKKTGSRPV	93
Db				
	1	VHMPKAVDPEAFNNISIIQHOGYPCPEEYEVATEDGYILSVNRI	PGLVOPKKTGSRPV	60
Qy	94	VLLQHGVLGGASWNISLNPNSLGLFIADAGDFVWNGSRGNAN	SRKHKTULSI	153
Db				
	61	VLLQHGVLGGASWNISLNPNSLGLFIADAGDFVWNGSRGNAN	SRKHKTULSI	120
Qy	154	FSYDEMARFDLPVINFILQKTGGEKIIYVYGVSOGTGMGFAT	STMPPELAQIK	213
Db				
	121	FSYDEMARFDLPVINFILQKTGGEKIIYVYGVSOGTGMGFAT	STMPPELAQIK	180
Qy	214	PIATVKHAKSPGTFFLLPDMNMIKGLFGKKEFLYQTRFL	RQLRVIIYLCQV	273
Db				
	181	PIATVKHAKSPGTFFLLPDMNMIKGLFGKKEFLYQTRFL	RQLRVIIYLCQV	240
Qy	274	LLLGGFNTNNMNSRASYAAHTLAGTSVQNILHWSQAVNS	GEURAFDWSG	333
Db				
	241	LLLGGFNTNNMNSRASYAAHTLAGTSVQNILHWSQAVNS	GEURAFDWSG	300
Qy	334	QPTPVRYRVFDMTVPYTAWMTGGQDWLSPNPDVKMLLS	EVNTNLIYHKNI	393

|||||  
Db 301 QTPVRYRDMTPTAMWTGQDLSNPEDVKMLLSEVTLNLYHKNIPEWAHVDFINGL 360  
QY 394 DAPHRMYNEIHLMOOEETNLSQGRCEAVL 423  
Db 361 DAPHRMYNEIHLMOOEETNLSQGRCEAVL 390

RESULT 4  
US-09-811-825-4  
; Sequence 4, Application US/09811825  
; Patent No. US20020144297A1  
; GENERAL INFORMATION:  
; APPLICANT: YAN, Chunhua et al.  
; TITLE OF INVENTION: ISOLATED HUMAN LIPASE PROTEINS, NUCLEIC  
; ACID MOLECULES ENCODING HUMAN LIPASE PROTEINS, AND USES  
; THEREOF  
; FILE REFERENCE: CLO01170  
; CURRENT APPLICATION NUMBER: US/09/811,825  
; CURRENT FILING DATE: 2001-03-20  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 374  
; TYPE: PRT  
; ORGANISM: Human  
US-09-811-825-4

Query Match 88.6%; Score 1991; DB 10; Length 374;  
Best Local Similarity 100.0%; Pred. No. 4.7e-180;  
Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 50 SEIIHQGYPCIEYEYVATEDGYILSVNRIPRGLVQPKTKGSRPVLLQHLGVLGGASNWIS 109  
Db 1 SEIIHQGYPCIEYEYVATEDGYILSVNRIPRGLVQPKTKGSRPVLLQHLGVLGGASNWIS 60  
QY 110 NLPNNSLGFILADAGFDVWNGSRGNAMSRKHKTLSIDODEFWAFSYDEMARFDPALVIN 169  
Db 61 NLPNNSLGFILADAGFDVWNGSRGNAMSRKHKTLSIDODEFWAFSYDEMARFDPALVIN 120  
QY 170 FILQKTGOEKIYVGVYSGQTGMGFIATFSTMPPELAQIKMYFALAPIATYKHAKEGPTKEL 229  
Db 121 FILQKTGOEKIYVGVYSGQTGMGFIATFSTMPPELAQIKMYFALAPIATYKHAKEGPTKEL 180  
QY 230 LLPDMIKGLFGKKEFLYQTRFLRQLVYLCGVILDOICSNIMLLGGFNTNNMNSRA 289  
Db 181 LLPDMIKGLFGKKEFLYQTRFLRQLVYLCGVILDOICSNIMLLGGFNTNNMNSRA 240  
QY 290 SVYAAHTLAGTSVONTLHWSQAVNSGELRAFDWSETKLEKNCQPTPVRYRDMTPT 349  
Db 241 SVYAAHTLAGTSVONTLHWSQAVNSGELRAFDWSETKLEKNCQPTPVRYRDMTPT 300  
QY 350 AMWTGGDLSNPEDVKMLLSEVTLNLYHKNIPEWAHVDFINGLDAPHMYNEIHLMOQ 409  
Db 301 AMWTGGDLSNPEDVKMLLSEVTLNLYHKNIPEWAHVDFINGLDAPHMYNEIHLMOQ 360  
QY 410 EETNLSQGRCEAVL 423  
Db 361 EETNLSQGRCEAVL 374

RESULT 5  
US-09-811-825-2  
; Sequence 2, Application US/09811825  
; Patent No. US20020144297A1  
; GENERAL INFORMATION:  
; APPLICANT: YAN, Chunhua et al.  
; TITLE OF INVENTION: ISOLATED HUMAN LIPASE PROTEINS, NUCLEIC  
; ACID MOLECULES ENCODING HUMAN LIPASE PROTEINS, AND USES  
; THEREOF  
; FILE REFERENCE: CLO01170  
; CURRENT APPLICATION NUMBER: US/09/811,825  
; CURRENT FILING DATE: 2001-03-20

; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 398  
; TYPE: PRT  
; ORGANISM: Human  
US-09-811-825-2

Query Match 88.6%; Score 1991; DB 10; Length 398;  
Best Local Similarity 100.0%; Pred. No. 5.1e-180;  
Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 50 SEIIHQGYPCIEYEYVATEDGYILSVNRIPRGLVQPKTKGSRPVLLQHLGVLGGASNWIS 109  
Db 25 SEIIHQGYPCIEYEYVATEDGYILSVNRIPRGLVQPKTKGSRPVLLQHLGVLGGASNWIS 84  
QY 110 NLPNNSLGFILADAGFDVWNGSRGNAMSRKHKTLSIDODEFWAFSYDEMARFDPALVIN 169  
Db 85 NLPNNSLGFILADAGFDVWNGSRGNAMSRKHKTLSIDODEFWAFSYDEMARFDPALVIN 144  
QY 170 FILQKTGOEKIYVGVYSGQTGMGFIATFSTMPPELAQIKMYFALAPIATYKHAKEGPTKEL 229  
Db 145 FILQKTGOEKIYVGVYSGQTGMGFIATFSTMPPELAQIKMYFALAPIATYKHAKEGPTKEL 204  
QY 230 LLPDMIKGLFGKKEFLYQTRFLRQLVYLCGVILDOICSNIMLLGGFNTNNMNSRA 289  
Db 205 LLPDMIKGLFGKKEFLYQTRFLRQLVYLCGVILDOICSNIMLLGGFNTNNMNSRA 264  
QY 290 SVYAAHTLAGTSVONTLHWSQAVNSGELRAFDWSETKLEKNCQPTPVRYRDMTPT 349  
Db 265 SVYAAHTLAGTSVONTLHWSQAVNSGELRAFDWSETKLEKNCQPTPVRYRDMTPT 324  
QY 350 AMWTGGDLSNPEDVKMLLSEVTLNLYHKNIPEWAHVDFINGLDAPHMYNEIHLMOQ 409  
Db 325 AMWTGGDLSNPEDVKMLLSEVTLNLYHKNIPEWAHVDFINGLDAPHMYNEIHLMOQ 384  
QY 410 EETNLSQGRCEAVL 423  
Db 385 EETNLSQGRCEAVL 398

RESULT 6  
US-10-042-431-41  
; Sequence 41, Application US/10042431  
; Publication No. US20020182675A1  
; GENERAL INFORMATION:  
; APPLICANT: MCCARTHY, Sean A  
; APPLICANT: BARNES, Thomas M  
; APPLICANT: FRASER, Christopher C  
; APPLICANT: SHARP, John D  
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,  
; PREVENTIVE, THERAPEUTIC, AND OTHER USES  
; FILE REFERENCE: 10147-602  
; CURRENT APPLICATION NUMBER: US/10/042,431  
; CURRENT FILING DATE: 2001-10-25  
; PRIOR APPLICATION NUMBER: US 09/333,159  
; PRIOR FILING DATE: 1999-06-14  
; PRIOR APPLICATION NUMBER: US 09/578,063  
; PRIOR FILING DATE: 2000-05-24  
; NUMBER OF SEQ ID NOS: 79  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 41  
; LENGTH: 399  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-042-431-41

Query Match 57.4%; Score 1289; DB 9; Length 399;  
Best Local Similarity 63.5%; Pred. No. 8.4e-114;  
Matches 235; Conservative 55; Mismatches 80; Indels 0; Gaps 0;  
QY 40 AVDPEAFMNISETIIOHGYPCIEYEYVATEDGYILSVNRIPRGLVQPKTKGSRPVLLQHLG 99  
||||| :||||| :||| :||||| :||| :||| :|||





```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 75
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-042-431-75

Query Match      49.8%; Score 1119; DB 9; Length 398;
Best Local Similarity 52.2%; Pred. No. 9.1e-98;
Matches 210; Conservative 73; Mismatches 107; Indels 12; Gaps 2;

Qy 17 MWLLTLVAYMFQNRVNSVHMTKAV-----DPEAFMNISEIIHQGYPCEEYEVATED 69
Db 1 MWLLTLVAYMFQNRVNSVHMTKAV-----DPEAFMNISEIIHQGYPCEEYEVATED 69

Qy 70 GYLSVNRIPRGLVQPKTGRPVVLLQHLGVLGGASNWSINLNNLSGLFILADAGFYVM 129
Db 56 GYILEVNRIPYKKNKSGNTGQRPVFLQHLGLASATNWSINLNNLSGLFILADAGFYVM 115

Qy 130 GNSRGNASRKHTLSIDODEFWAFSYDEMARFDLPVAVINFLQKTGQEKIYYVGSQGT 189
Db 116 GNSRGNTWARRNLYTSPDSVEFWAFSDEMAKYDLPATIDFVKKTKGQKQLHYVGSQGT 175

Qy 190 TWGFTAFSTMPELAQIKMYFALAPIATVKHAKSPCTKFLLLPDMWIKGLGKKEFLYQT 249
Db 176 TWGFTAFSTMPELAQIKMYFALAPIATVKHAKSPCTKFLLLPDMWIKGLGKKEFLYQT 235

Qy 250 RFLRLVLYLCOVLDQICSNIMLLGGFNTNNMNSRASVYAAHTLAGTSVQNILHWS 309
Db 236 RFLRLVLYLCOVLDQICSNIMLLGGFNTNNMNSRASVYAAHTLAGTSVQNILHWS 295

Qy 310 QAVNSGELEAFDWGSETKLEKCNQPTPVRYRVRDMTPTAMWTGGQDMLSNPDKMLL 369
Db 296 QAVNSGELEAFDWGSETKLEKCNQPTPVRYRVRDMTPTAMWTGGQDMLSNPDKMLL 355

Qy 370 SEVTNLYHKNIPEVAHVDFIWLGLDAPHRMYNEIIHLMOQEE 411
Db 356 SEVTNLYHKNIPEVAHVDFIWLGLDAPHRMYNEIIHLMOQEE 397

RESULT 10
US-10-056-744B-2
; Sequence 2, Application US/10056744B
; Publication No. US2002019303A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: 58860, A HUMAN CHOLESTERYL ESTER
; TITLE OF INVENTION: HYDROLASE AND USES THEREFOR
; FILE REFERENCE: MPI2001-026PIRNM
; CURRENT APPLICATION NUMBER: US/10/056.744B
; PRIOR FILING DATE: 2002-08-16
; PRIOR FILING DATE: 60/264,167
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-056-744B-2

Query Match      48.5%; Score 1090; DB 9; Length 398;
Best Local Similarity 52.4%; Pred. No. 5e-95;
Matches 207; Conservative 61; Mismatches 123; Indels 4; Gaps 2;

Qy 17 MWLLTLVAYMFQNRVNSVHMTKAVDPEAFMNISEIIHQGYPCEEYEVATEDGYIL 73
Db 2 MWLLTLVAYMFQNRVNSVHMTKAVDPEAFMNISEIIHQGYPCEEYEVATEDGYIL 60

Qy 74 SVNRIPRGLVQPKTGRPVVLLQHLGVLGGASNWSINLNNLSGLFILADAGFYVMGNSR 133
Db 61 LVNRIPYGRTHARSTGPRPVVYMQHALFADNAYWLENYANGSLGFLADAGFYVMGNSR 120

Qy 134 GNSRGNASRKHTLSIDODEFWAFSYDEMARFDLPVAVINFLQKTGQEKIYYVGSQGT 193
Db 121 GNSRGNASRKHTLSIDODEFWAFSYDEMARFDLPVAVINFLQKTGQEKIYYVGSQGT 180

Qy 194 TAFSTMPELAQIKMYFALAPIATVKHAKSPCTKFLLLPDMWIKGLGKKEFLYQTRFLR 253
Db 181 TAFSTMPELAQIKMYFALAPIATVKHAKSPCTKFLLLPDMWIKGLGKKEFLYQTRFLR 240

Qy 254 QLVIVLYLCOVLDQICSNIMLLGGFNTNNMNSRASVYAAHTLAGTSVQNILHWSQAVN 313
Db 241 TASTKICNKKILWLICSEFMSLWAGSNKKNMQSRMDVYMSHAPTSSVHNILHIKQLYH 300

Qy 314 SGELRAFDMGSETKLEKCNQPTPVRYRVRDMTPTAMWTGGQDMLSNPDKMLLSEVT 373
Db 301 SDEFRAYDNGDADNKKHYNQSHPPYIDLTAMKVPTAIWAGGHVLTVPQDVARILPQIK 360

Qy 374 NL-IYHKNIPEVAHVDFIWLGLDAPHRMYNEIIHLMO 401
Db 361 SLSLVLTLPEWEPTDFVWGLDAPQRMFS 390
```

Query Match 5.7%; Score 128; DB 9; Length 25;  
Best Local Similarity 100.0%; Pred. NO. 5.5e-06;  
Matches 25; Conservative 0; Mismatches 0; Indels

```
Query Match      17.8%; Score 400; DB 9; Length 127;
Best Local Similarity   56.5%; Pred. NO. 1.2e-30;
Matches 72; Conservative 22; Mismatches 29; Indels    0; Gaps
```

Oy 107 WISLNPNSLGFIILADAGFDVWMSNRSGNAWSRRKHTLSIDQDFWFASFYDEMAREFDLP A 166  
1 WIANGPKSLAFIILADAGDYDVLNGSNGNTYSRKKVKLNPSHSEFFWDFSWHGMYDLPA 60  
Db

Search completed: February 3, 2003, 13:54:11  
Job time : 14 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:47:54 ; Search time 21 Seconds  
(without alignments)  
1936.422 Million cell updates/sec

Title: US-09-333-159-47

Perfect score: 2247

Sequence: 1 MLETLRQWIVSRHMEMWLL.....IHLMOQETNLSQRCCEAVL 423

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR73.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1290	57.4	399	2 S41408	lysosomal acid lip
2	1290	57.4	399	2 G01416	lysosomal acid lip
3	1164.5	51.8	395	1 L1RTT	triacylglycerol li
4	1139	50.7	397	1 JC4017	triacylglycerol li
5	1119	49.8	398	2 S07145	triacylglycerol li
6	752.5	33.5	403	2 T33198	hypothetical prote
7	751	33.4	411	2 T22290	hypothetical prote
8	730	32.5	405	2 T22675	hypothetical prote
9	707.5	31.5	405	2 H88930	protein R11G1.14
10	692	30.8	411	2 C89074	protein K0488.5 [1
11	655.5	29.2	426	2 T20480	hypothetical prote
12	574.5	25.6	559	2 JT0949	egg-specific prote
13	505	22.5	443	2 T39540	triacylglyceride lipas
14	478.5	21.3	1585	2 T31611	hypothetical prote
15	475	21.1	467	2 T41053	triacylglyceride lipas
16	445	19.8	344	2 E84526	probable lysosomal
17	422.5	18.8	548	2 S37969	probable triacylgl
18	412.5	18.4	460	2 T39443	probable triglycer
19	411.5	18.3	413	2 T43170	probable triacylgl
20	294	13.1	431	2 D86318	protein F15H18.6 [
21	285	12.7	538	2 S64842	probable membrane
22	271.5	12.1	573	2 S64754	probable membrane
23	234.5	10.4	509	2 G96766	protein lipase F2P
24	149	6.6	62	2 S59904	lipase, pregastric
25	143	6.4	336	2 F83425	probable esterase/
26	126	5.6	987	2 A64474	hypothetical prote
27	113	5.0	412	2 B72391	conserved hypothet
28	112.5	5.0	40	2 S19539	triacylglycerol li
29	107	4.8	281	2 A11912	lysophospholipase

ALIGNMENTS

RESULT 1

S41408

lysosomal acid lipase (EC 3.1.1.1.-) / sterol esterase (EC 3.1.1.13) precursor - hum  
C:Species: Homo sapiens (man)  
C:Date: 13-Jan-1995 #sequence-revision 13-Jan-1995 #text-change 18-Jun-1999  
C:Accession: S41408; A39315; S47187  
R:Ames, D.; Merkel, M.; Eckerskorn, C.; Greten, H.  
Eur. J. Biochem. 219, 905-914, 1994  
A:Title: Purification, characterization and molecular cloning of human hepatic lys.  
A:Reference number: S41408; MUID:94155897; PMID:8112342  
A:Accession: S41408  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-399 <AME>  
A:Cross-references: EMBL:X76488; NID:9434305; PIDN:CAA54026.1; PID:9434306  
J. Biol. Chem. 265, 22479-22484, 1991  
A:Title: Cloning and expression of cDNA encoding human lysosomal acid lipase/chole  
A:Reference number: A39315; MUID:92042192; PMID:1718995  
A:Accession: A39315  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-15, 'P', 17-399 <AND>  
A:Cross-references: GB:M74775; NID:9187151; PIDN:AAA59519.1; PID:9187152  
R:Du, H.; Gregory, G.A.  
submitted to the EMBL Data Library, April 1994  
A:Description: Structural conservation of putative functional motifs between mouse

hypothetical prote  
hypothetical prote  
probable protease  
OatR protein - Eme  
platelet-activatin  
probable 2-acetyl-  
hypothetical prote  
probable permease  
ABC transporter, m  
unknown protein F2  
probable epha prot  
hypothetical prote  
hypothetical prote  
reverse transcript  
probable membrane  
DNA polymerase I V

Query Match 57.4%; Score 1290; DB 2; Length 399;  
Best Local Similarity 61.0%; Pred. No. 1.8e-105;  
Matches 242; Conservative 59; Mismatches 94; Indels 2; Gaps 1;

Qy 15 MENWLLIYVAFQFNVSVMPTK--AVDPEAFNISEIIHQHQPCEYEVAEDGYI 72

Db 1 MKNRFGLVGLVCLVTLHSESGGKLTAVDPETNNVSEIISYWGFPPEEYLVETEDGYI 60

Qy 73 LSVNRIPLGLVQPKTGRPVVLLQHLGVGGASNIPLNPNNSLGLADAGFDVWMGNS 132

Db 61 LCLANRPHGRKNHSDKPKVPVFLQRLADSSNVTNLANSLSGLADAGFDVWMGNS 120

QY 133 RGNWSRKHKTLSDQDEFWAFSYDEMAREDLPAVINFILOKTGOEKIYVYVSGQTTMG 192  
Db 121 RGNWSRKHKTLSDQDEFWAFSYDEMAREDLPAVINFILOKTGOEKIYVYVSGQTTIG 180  
QY 193 FIAPSTPELAQIKMYFALAPATVTKHAKSPGTFKLLPDMIMKGLFGKKEFLYQTRFL 252  
Db 181 FIAPSQIPELAKRIMKMFALGPVAVAFCSFPAKGLRGLPDHLIKDLFGDKEFLPQSAFL 240  
QY 253 ROLVYLCGVILDOICSNMILLGGFTNTNMNMSRASVYAHTLAGTSVQNLHWSQAV 312  
Db 241 KWLGTHTVCTHVKELCGNLCFLCCGFEENLNMSRDVVTTHSPAGTSVQNLHWSQAV 300  
QY 313 NSGELRAFDMGSETKNEKCNQPTPVRYRDMTPVAMTGGODWLSNPDYKMLLSEV 372  
Db 301 KFOKFOAFDGMSSAKNFHYNQSPPTTYNKDMLVPTAVMSGGHDLADYVDVNIILTOI 360  
QY 373 TNLVYHKNIPEWAHVDFIWLGLDAPHRYNEIHLMOQ 409  
Db 361 TNLVYHESIPWEHLDFIWLGLDAPHRYNEIHLMOQ 397  
RESULT 2  
G01416  
lysosomal acid lipase - human  
C:Species: Homo sapiens (man)  
C:Date: 21-Dec-1996 #sequence\_revision 06-Jun-1997 #text\_change 18-Jun-1999  
C:Accession: G01416  
R:Du, H.  
submitted to the EMBL Data Library, April 1994  
A:Reference number: G06919  
A:Accession: G01416  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-399 <DUX>  
A:Cross-references: EMBL:U08464; NID:G505052; PIDN:AAB60328.1; PID:G505053  
C:Superfamily: triacylglycerol lipase, lingual  
C:Keywords: glycoprotein  
Query Match 57.4%; Score 1290; DB 2; Length 399;  
Best Local Similarity 61.0%; Pred. No. 1.8e-105;  
Matches 242; Conservative 59; Mismatches 94; Indels 2; Gaps 1;  
QY 15 MEMLLILVAYMFORNVNVMPTK--AVDPEAFMNISEIIHQGYPCPEEYEVATEDGYI 72  
Db 1 MKRFLGLVCLVLTWTHSEGRKGLTAVDPETNNMYSEIISYNGFPSEELYVEDGYI 60  
QY 73 LSVNRIPRGLVQPKTSRPVLLQHLGVLGASNWSINLNNLSGLFLADAGFDVWNGNS 132  
Db 61 LCLNRIPIHGRKNNHSDGKPKPVYFLQHLGLADSSNNWVNLANSLSGLFLADAGFDVWNGNS 120  
QY 133 RGNWSRKHKTLSDQDEFWAFSYDEMAREDLPAVINFILOKTGOEKIYVYVSGQTTMG 192  
Db 121 RGNWSRKHKTLSDQDEFWAFSYDEMAREDLPAVINFILOKTGOEKIYVYVSGQTTIG 180  
QY 193 FIAPSTPELAQIKMYFALAPATVTKHAKSPGTFKLLPDMIMKGLFGKKEFLYQTRFL 252  
Db 181 FIAPSQIPELAKRIMKMFALGPVAVAFCSFPAKGLRGLPDHLIKDLFGDKEFLPQSAFL 240  
QY 253 ROLVYLCGVILDOICSNMILLGGFTNTNMNMSRASVYAHTLAGTSVQNLHWSQAV 312  
Db 241 KWLGTHTVCTHVKELCGNLCFLCCGFEENLNMSRDVVTTHSPAGTSVQNLHWSQAV 300  
QY 313 NSGELRAFDMGSETKNEKCNQPTPVRYRDMTPVAMTGGODWLSNPDYKMLLSEV 372  
Db 301 KFOKFOAFDGMSSAKNFHYNQSPPTTYNKDMLVPTAVMSGGHDLADYVDVNIILTOI 360  
QY 373 TNLVYHKNIPEWAHVDFIWLGLDAPHRYNEIHLMOQ 409  
Db 361 TNLVYHESIPWEHLDFIWLGLDAPHRYNEIHLMOQ 397

RESULT 3  
LIRTT

triacylglycerol lipase (EC 3.1.1.3) precursor, lingual - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 30-Sep-1987 #sequence\_revision 30-Sep-1987 #text\_change 11-Jun-1999  
C:Accession: A23045  
R:Doeherty, A.J.P.; Bodmer, M.W.; Angal, S.; Verger, R.; Riviere, C.; Love, P.A.; Lyo  
Nucleic Acids Res. 13, 1891-1903, 1985  
A:Title: Molecular cloning and nucleotide sequence of rat lingual lipase cDNA.  
A:Reference number: A23045; MUID:85215587; PMID:3839077  
A:Accession: A23045  
A:Molecule type: mRNA  
A:Residues: 1-395 <DUC>  
A:Cross-references: GB:X02309; NID:G56595; PIDN:CAA26179.1; PID:G56596  
A:Experimental source: strain Sprague-Dawley  
A:Note: The partial sequence of the mature protein from a different, unspecified stra  
C:Comment: This acid-stable lipase is secreted by the serous (von Ebner's) glands at  
C:Superfamily: triacylglycerol lipase, lingual  
C:Keywords: carboxylic ester hydrolase; glycoprotein; lipid digestion; saliva; serous  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>  
F:33,68,98,184,270/Binding site: carbohydrate (Asn) (covalent) #status Predicted  
Query Match 51.8%; Score 1164.5; DB 1; Length 395;  
Best Local Similarity 54.9%; Pred. No. 1.9e-94;  
Matches 217; Conservative 69; Mismatches 108; Indels 1; Gaps 1;  
QY 17 MWLLILVAYNFQ-RNVNSVHMPKAVDPEAFMNISEIIHQGYPCPEEYEVATEDGYILSV 75  
Db 1 MWLLITSVISTFGAGHGLFKLPGCPNPEANMISQMITWYPCQYEVVEDGYILGV 60  
QY 76 NRIPRGLVQPKTSRPVLLQHLGVLGASNWSINLNNLSGLFLADAGFDVWNGNSRGN 135  
Db 61 YRIPHGKNNSEINIGKRPVYVQLHGLIASATNWTIANLNNLSLAFLADAGFDVWNGNSRGN 120  
QY 136 AWSRKHKTLSDQDEFWAFSYDEMAREDLPAVINFILOKTGOEKIYVYVSGQTTMGFIA 195  
Db 121 TWSRKNVYSPDSVEFWAFSDENAKYDLPATINFIYQKTGOEKIHYVSGQTTIGFIA 180  
QY 196 FSTPELAQIKMYFALAPATVTKHAKSPGTFKLLPDMIMKGLFGKKEFLYQTRFLRQL 255  
Db 181 FSTNPTLAKIKTFYALAPATVTKYQSPKLSIFITFLFKLMGKMKFLPHTYFDDEL 240  
QY 256 VIYCGVILDOICSNMILLGGFTNTNMNMSRASVYAHTLAGTSVQNLHWSQAVNSG 315  
Db 241 GTEVCSREVLIDLCNTLFIQCGDKKLNLYSRPDVYLGHNPAGTSVQDLHMAQLVRS 300  
QY 316 ELRAFDMGSETKNEKCNQPTPVRYRDMTPVAMTGGODWLSNPDYKMLLSEVTLN 375  
Db 301 KFOAFNNGSFSQNLHYNQKTPPEYDVSAMTPVAVVNGNDILADPDVAMLLPKLSNL 360  
QY 376 IYHKNIPWEAHDVFIWLGLDAPHRYNEIHLMOQ 410  
Db 361 LFHKEILAYNHLDFIWLADAPQEVYNEIMSMAD 395  
RESULT 4  
JC4017  
triacylglycerol lipase (EC 3.1.1.3) PGE precursor - bovine  
N:Alternate names: pregastric esterase  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: JC4017; S64678  
R:Timmermans, M.Y.J.; Teuchy, H.; Kupers, L.P.M.  
Gene 147, 259-262, 1994  
A:Title: The cDNA sequence encoding bovine pregastric esterase.  
A:Reference number: JC4017; MUID:95011625; PMID:7926811  
A:Accession: JC4017  
A:Molecule type: mRNA  
A:Residues: 1-397 <TIM>  
A:Cross-references: GB:L26319; NID:G600756; PIDN:AAA57037.1; PID:G600757  
A:Experimental source: tongue  
R:Timmermans, M.Y.J.; Reekmans, G.; Teuchy, H.J.H.; Kupers, L.P.M.  
Biochem. J. 314, 931-936, 1996  
A:Title: Inhibition studies on calf pregastric esterase: the enzyme has no functional

A:Reference number: S64678; MUID:96177869; PMID:8615791

A:Accession: S64678

A:Molecule type: protein

A:Residues: 24-40;248-253 <NID>

C:Comment: Pregastric esterase is a major fat-digesting enzyme.

C:Genetics:

A:Gene: pge

C:Superfamily: triacylglycerol lipase, lingual

C:Keywords: blocked amino end; carboxylic ester hydrolase; glycoprotein

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-397/Product: pregastric esterase #status predicted <MAT>

F:33,270,326/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:171/Active site: Ser #status predicted

Query Match 50.7%; Score 1139; DB 1; Length 397;  
Best Local Similarity 53.7%; Pred. NO. 3.4e-92;  
Matches 212; Conservative 68; Mismatches 113; Indels 2; Gaps 2;

QY 18 WLLILVAYMFQQRVNSVHPTKAVDPEAFNMNISEIIHQGYPCPEYEVATEDGYILSVNR 77

Db 3 WLLVTVCFTHMGNACFLGKIAKNPEASMNYSOMISYMGYPSEMHKVTADGYILOVYR 62

QY 78 IPRGLVQPKTKSRPVVLLQHLGVLGASNWSNLNPNNSLGFILADAGDVNMGNSRGNW 137

Db 63 IPRGNANHLQRPVVFQHLGLSATNWSNLNPNNSLGFILADAGDVNMGNSRGNW 122

QY 138 SRKKHTLSIDODEWAFSYDEMAREFDLPVAVINFILOKTGOEKIYVVGSGTGMGFIAPS 197

Db 123 AQEHLVYSPDSPEFWAFSDEMAEYDLSTIDFLRRGOKKLHYVGHSGTGTGFIAPS 182

QY 198 TPELAOKTKMYFALAPIATVYKHAQSPGKFLLLPDMIMKGLFGKKEFLYOTRFRLQ-IV 256

Db 183 TSPTEAEKIKFYALAPVATVYKSLFNKLAIPHFLFKIIFGDKMF-YPHTELEQFLG 241

QY 257 IYLCGOVILDOCSNMLLLGFTNTNNMNSRASVYAAHTLAGTSVQNLHWSQAVNSGE 316

Db 242 VECRSRETLVDVLCCKNALFAITGVDNKNFNSRLDYIAHNPAGTSVQNLHWRQAVKSGK 301

QY 317 LRAFDGSGETKLNKCNQPTPVYRVDRMTVPTAMWTGGQDWSLPEDVKMLLSEVTNLI 376

Db 302 FOAFDGAQPYQNLHMYHQPTPIYNTLANNVPIAVNSADNLLADPDQVDVFLSKLSNLI 361

QY 377 YKKNIPENAHVDFIWLGLDAPHRYNEIHLMOQEE 411

Db 362 YKKEIPNHNHLDPIWADAPQEVYNEIVSLMAEDK 396

#### RESULT 5

triacylglycerol lipase (EC 3.1.1.3) precursor, gastric - human

C:Species: Homo sapiens (man)

C:Date: 29-Jan-1993 #sequence revision 29-Jan-1993 #text\_change 18-Jun-1999

C:Accession: S07145; S27102; S04942

R:Bottom: M.W.; Angal, S.; Varranton, G.T.; Harris, T.J.R.; Lyons, A.; King, D.J.; Ptero

Biochem. Biophys. Acta 909, 237-244, 1987

A:Title: Molecular cloning of a human gastric lipase and expression of the enzyme in yea

A:Reference number: S07145; MUID:87299724; PMID:3304425

A:Accession: S07145

A:Molecule type: mRNA

A:Residues: 1-398 <BODI>

A:Cross-references: EMBL:X05997; NID:g31771; PIDN:CAA29413.1; PID:g758063

A:Accession: S27102

A:Molecule type: protein

A:Residues: 20-24,'X',26,'X',28-33,'X',35-42,'X',44,'X',46-48,'X',50,'X',52,'X',54-56 <E

A:Note: It is uncertain whether Met-1 or Met-7 is the initiator

R:Bernbaeck, S.; Blaeckberg, L.

Eur. J. Biochem. 182, 495-499, 1989

A:Reference number: S04942; MUID:89325292; PMID:2753032

A:Accession: S04942

A:Molecule type: protein

A:Residues: 20-45 <BER>

C:Superfamily: triacylglycerol lipase, lingual

C:Keywords: carboxylic ester hydrolase; extracellular protein; glycoprotein; lipid  
F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-398/Product: triacylglycerol lipase, gastric #status experimental <MAT>

F:34,99,185,271/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 49.8%; Score 1119; DB 2; Length 398;

Best Local Similarity 52.2%; Pred. NO. 2e-90;

Matches 210; Conservative 73; Mismatches 107; Indels 12; Gaps 2;

QY 17 MWLLLVAYMFQQRVNSVHPTKAV-----DPEAFNMNISEIIHQGYPCPEYEVATED 69

Db 1 MWLLLVAYMFQQRVNSVHPTKAV-----SVLGTHGLFGLKHPGSPVMTNISMIIYWGYPNEEYVATED 55

QY 70 GYILSVNRIPRGLVQPKTKSRPVVLLQHLGVLGASNWSNLNPNNSLGFILADAGDVNWM 129

Db 56 GYILEVNRIPYKKNKSGTQRPVVFQHLGLSATNWSNLNPNNSLGFILADAGDVNWL 115

QY 130 GNSRGNASRKHHTLSIDODEWAFSYDEMAREFDLPVAVINFILOKTGOEKIYVVGSGT 189

Db 116 GNSRGNWARRNLYYSPDSPEFWAFSDEMAKYDLPATIDFIVKTKGQKQLHYVGHSGT 175

QY 190 TMGTFATSTMPLELAOKIKMYFALAPIATVYKHAQSPGKFLLLPDMIMKGLFGKKEFLYOT 249

Db 176 TIGTFATSTNPSLARIKTFYALAPVATVYKSLINKLRFVPSQSLFKIEFGDKIEYPHN 235

QY 250 RFLRLVLYLCGOVILDOCSNMLLLGFTNTNNMNSRASVYAAHTLAGTSVQNLHWS 309

Db 236 FFIQPLATEVCSREMLNLLCSNALFICGFSKNFTSLDVLVLSHNPAGTSVQNMFWHT 295

QY 310 QAVNSGELRAFQWGSQETKLNKCNQPTPVYRVDRMTVPTAMWTGGQDWSLPEDVKMLL 369

Db 296 QAVKSGKQAYDWGSPVQNRHMDQSPPYNYNTAMNVPVAVNMGKDLADPDQVGLLL 355

QY 370 SEVTNLIYKKNIPENAHVDFIWLGLDAPHRYNEIHLMOQEE 411

Db 356 PKLPNLIIHKIEPFYNHLLDFIWMADAPQEVYNDIVSMISEDK 397

#### RESULT 6

T33198

hypothetical protein ZK6.7 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence revision 29-Oct-1999 #text\_change 21-Jan-2000

C:Accession: T33198

R:Wu, X.

submitted to the EMBL Data Library, May 1998

A:Description: The sequence of C. elegans cosmid ZK6.

A:Reference number: Z21301

A:Accession: T33198

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-403 <WUX>

A:Cross-references: EMBL:AF067942; PIDN:AAC17694.1; GSPDB:GN00023; CESP:ZK6.7

A:Experimental source: strain Bristol N2; clone ZK6

C:Genetics:

A:Gene: CESP:ZK6.7

A:Map position: 5

C:Superfamily: triacylglycerol lipase, lingual

Query Match 33.5%; Score 752.5; DB 2; Length 403;  
Best Local Similarity 41.1%; Pred. NO. 3.5e-58;  
Matches 168; Conservative 66; Mismatches 142; Indels 33; Gaps 12;

QY 17 MW-LLILVAYMFQQRVNSVHPTKAVDPEAFNMNISEIIHQGYPCPEYEVATEDGYILSV 75

Db 1 MRFVAVFLAAFFVGVGSH-----GDPELHMTTPQIERWGYPMIYVATDDGYILEM 55

QY 76 NRTPRGLVQPKTK-----GSRPVVLLQHLGVLGASNWSNLNPNNSLGFILADAGDVNWM 129

Db 56 HRIIPFG-----KTNVTWPNKGRPVVFNQGLLCASDWWVNLFPDQAGSGLFADAGFDVWL 110

QY 130 GNSRGNASRKHHTLSIDODEWAFSYDEMAREFDLPVAVINFILOKTGOEKIYVVGSGT 189

Db 111 GNARGNTSMKHKDLKPSHAFDWSDEMATYDLNAMINHVLVETGQDSVYVYMGHSQGT 170  
QY 190 TMGFIAPSTWP-ELAAQIKMYFALAPATYKHAQSPGTEKLLLPDMNMGK--LFGKKEF 245  
Db 171 LWFHSLSKDGSFAKAKKFFALAPIGSVKHKIGFLSFFANFYSLEFDGWFDFGAGEF 230  
QY 246 LYQTRFLRQVILVCGOVIDL-QICSNMILLGFGFNTNNMNSRASVYAAHTLAGTSVQN 304  
Db 231 LPNNWAKLAADICGKGLVEADLCNDVFLIAGPESDQWNOVTRPVYATHDPAGTSTQN 290  
QY 305 ILHWSQAVNSGELRAFDMGSETKNLEKCNQPTPVRYVRDMT-----VPTAMWTGGQDWLS 360  
Db 291 IVHWMQVHGGVPAYDMGKT-NKKYGOANPEY---DFTAIKGTKIYLYSDADWLA 346  
QY 361 NPEDV-KMLLSEVTLNLYHKN-IPEWAHVDFTWGLDAPHRMYNEIHL 406  
Db 347 DTPDVPDYLTRLNPAIVAOHNLPLDYNHLDFTWGLRAPODIYRPAIKL 395

RESULT 7  
T22290  
hypothetical protein F46B6.8 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 21-Jan-2000  
C:Accession: T22290  
R:McMurray, A.  
submitted to the EMBL Data Library, April 1996  
A:Reference number: Z19542  
A:Accession: T22290  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-411 <N1>  
A:Cross-references: EMBL:Z70780; PIDN:CAA94824.1; GSPDB:GN00023; CESP:F46B6.8  
A:Experimental source: clone F46B6  
C:Genetics:  
A:Gene: CESP:F46B6.8  
A:Map position: 5  
A:Introns: 35/3; 69/1; 104/1; 226/3  
C:Superfamily: triacylglycerol lipase, lingual

Query Match 33.4%; Score 751; DB 2; Length 411;  
Best Local Similarity 40.4%; Pred. No. 4.9e-58;  
Matches 163; Conservative 72; Mismatches 138; Indels 30; Gaps 12;  
QY 20 LILVAYFORNVSVHMTKAVDPEAFNISEIIHQHGYPCPEEYEVATEDGYILSVNRIP 79  
Db 10 LICTSFLLOTINA-HPDD---DPELNNTSQIIRGWYKAEVHTVTTEDGYILOMQRIP 65  
QY 80 RGLVQPKKT-----GSRPVVLLQRLVGGASNWSLNNPNSLGFILADAGFDVMMGNSR 133  
Db 66 YG-----KTSVTPWNGKRPVILLQHLGACASDWDNLPQTQSAFVADAGFDVNLGNVR 120  
QY 134 GNWSRKHKTLSDIDQDFWAFSYDEMARFDLPVINFILQKTGQEKIYVYVSGTGMGF 193  
Db 121 GTTYGRKNTKLDPSSETAFQFSDWMAQYDVPAMVDVHVLAMTGOENLYYMGHSQCTIME 180  
QY 194 --IAFSTWPELAQIKMYFALAPATYKHAQSPGTEKLLLPDMNMGK--LFGKKEFIQ 248  
Db 181 THLAKTDGSGFAKKIKRYFALAPAGVKNITKGLFSLYFAHKFSPFEDGMWYELSKDFLPD 240  
QY 249 TRFLRQLVILVCGOVIDL-QICSNMILLGFGFNTNNMNSRASVYAAHTLAGTSVQNILH 307  
Db 241 NWITKMAKDICGASEAEKELCDNELFLIGPESDQWNSATYSSODPAGTSTQNIHV 300  
QY 308 WSAVNSGELRAFDMGSETKNLEKCNQPTPVRY---RVDMVTVPMTAMTGGQDWLSNPED 364  
Db 301 WQWVRNGRVPAPFDWGGKI-NKKYGOQDTPPEYDFGAIKGTKI--HLWSDDDMLGSDPTD 357  
QY 365 V-KMLLSEVTLNLYHKN-IPEWAHVDFTWGLDAPHRMYNEIHL 404  
Db 358 IHDFELKELNPAVIAENVLNKEYNHLDFTWGLNATFQIYDKAI 400

RESULT 8  
T22675  
hypothetical protein F54F3.3 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 21-Jan-2000  
C:Accession: T22675  
R:Percy, C.; Lloyd, C.  
submitted to the EMBL Data Library, September 1996  
A:Reference number: Z19598  
A:Accession: T22675  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-405 <N1>  
A:Cross-references: EMBL:Z79696; PIDN:CAB01973.1; GSPDB:GN00023; CESP:F54F3.3  
A:Experimental source: clone F54F3  
C:Genetics:  
A:Gene: CESP:F54F3.3  
A:Map position: 5  
A:Introns: 31/3; 65/1; 274/3; 366/3  
C:Superfamily: triacylglycerol lipase, lingual

Query Match 32.5%; Score 730; DB 2; Length 405;  
Best Local Similarity 38.6%; Pred. No. 3.4e-56;  
Matches 161; Conservative 69; Mismatches 155; Indels 32; Gaps 11;  
QY 15 MEMVLLILVAYMFQNVSVHMTKAVDPEAFNISEIIHQHGYPCPEEYEVATEDGYILS 74  
Db 1 MRSWSTVNLAV/LATAATVFGH---DADPEKMTPTQIIMRWGYPAMIYDVTEDGVILE 56  
QY 75 VNRIPRGLVQPKKT-----GSRPVVLLQHLVGGASNWSLNNPNSLGFILADAGFDVW 128  
Db 57 LHRIPYG-----KTNVTPWNGKRPVFMQHGLESCSSWVNVNLPESAAFLADAGYDVW 111  
QY 129 MGNRGNANWSEKHKTLSDIDQDFWAFSYDEMARFDLPVINFILQKTGQEKIYVYVSG 188  
Db 112 LGNFRGNTYSMKHKLPKSPHSAFMDSWDEMQQYDLPAMIEKALEVTSGLSYIYHSG 171  
QY 189 T-TMGFIATFSTPE---LAQIKMYFALAPATYKHAQSPGTEKLLLPDMNMGK---LF 240  
Db 172 TLTW---FSRLSEDKVGVGNKIKKFFALAPVGSVKIKKALKAFFADYFLEDFGDFV 227  
QY 241 GKKEFLYQTRFLRQLVILVCGOVIDLQICSNMILLGFGFNTNNMNSRASVYAAHTLAG 299  
Db 228 GSGEFLNNIMKLVSSEVAGLVKVEAGVDDVDFVLIAGPESNOLNATRPVYVAHTPAG 287  
QY 300 TSVONILHWSQAVNSGELRAFDMGSETKNLEKCNQPTPVRYVRDMTVPMTAMTGGQDWL 359  
Db 288 TSTONIVHWIWMVRHGGTPKYDYG-ERGNKKHYQANVPAYDFTTVNRVYLYWGDSDWL 346  
QY 360 SNPEDV-KMLLSEV--TNLYHKNIPESWAHVDFTWGLDAPHRMYNEIHLMOOETN 413  
Db 347 ADPTDVTDFLTHLPSTVQVNNKLIDYNHLDFTWGLRAPDIPEIPIIDIVRNDVLN 403

RESULT 9  
H88930  
protein R1LG11.14 [imported] - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 24-May-2001  
C:Accession: H88930  
R:anonymous, The C. elegans Sequencing Consortium.  
Science 282, 2012-2018, 1998  
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating bio  
A:Reference number: A75000; MUID:99089613; PMID:9851916  
A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C-  
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999;  
A:Accession: H88930  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-405 <STO>  
A:Cross-references: GB:chr.V; PIDN:AAC69088.1; PID:g2384863; GSPDB:GN00023; CESP:R1LG  
A:Note: Similar to lipase; R1LG11.14



C:Genetics:

A:Gene: R1LG11.14

A:Map position: 5

C:Superfamily: triacylglycerol lipase, lingual

Query Match 31.5%; Score 707.5; DB 2; Length 405;

Best Local Similarity 39.0%; Pred. No. 3.2e-54;

Matches 162; Conservative 75; Mismatches 133; Indels 45; Gaps 15;

QY 19 LLILVAYMFORNVNSVHMTKAVDPEAFNMNISEIIIOHGVPCPEYEVATEDGVILSVNRI 78

DB 9 LLVILA-----VHNVHAKS---DPELHMTTPOIIRWGYPMIYVTTDDGILDEHRI 59

QY 79 PRGLVOPKKT-----GSRPWVLLQHLGVGASNMISNLPNNSLGFILADAGDVWNGNS 132

DB 60 PHG-----KTNVTPNGKQPVPMQHLGCLASTDWTMNLPEQSAAFADAGDVWLGNN 114

QY 133 RGNWRSKHKHTLSIDODEWAFSYDEMAREFDLPAVINFLQKTQGEKIYVGVSGQT-TW 191

DB 115 RGNTYSKKHKNLAKASHDFEWSDEMATYDLPAMDKVLEVTGQESLYMGHSQGTLM 174

QY 192 GFIAFTMPD-----LAQIKMYFALAPIATVKHAKSPGKTFLLPDMIMKG---LFGKKE 244

DB 175 -----FHLSKDDGIFAKKIKKFFALAPVGSVKDIKFLGFHFFSLEFDGWDVFGAGE 230

QY 245 FLYQTRFLRQLVYILGQV--ILDQICSNIMLLGGFNMMNSRASVYAAHTLAGTSVQ 303

DB 231 FLPNWAMKLAADICGLKIESDLCDNVCFLIAGPESDOWNSTRVPVYASHDPAGTATQ 290

QY 304 NILHWSQVNSGELRAFDMGSETKNLEKCN-OPTPVRYVRVDMT-----VPTAMWTGGQDW 358

DB 291 NIVHWIQMVHGGVPAYDMGSK-ENKKNVNFOANPEY---DFTAIGTQIYLYWSADW 346

QY 359 LSNPEDEV-KMLLSEVNLVYHKN--IPEWAHVDFIAGLDAPHRYMNEIHLMOQEE 410

DB 347 LADKTDITNLLRLNPAIIAQNYYFTDYNHDFVFLRAPNDIYLPVDICTKD 401

RESULT 10

G89074

protein K04A8.5 [imported] - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 24-May-2001

C:Accession: G89074

R:Anonymous, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998

A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology

A:Reference number: A75000; MUID:99069613; PMID:9851916

A:Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/Projects/C\_ele

A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and

A:Accession: G89074

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-411 &lt;STO&gt;

A:Cross-references: GB:chr\_V; PIDN:AAC48051.1; PID:g1658354; GSPDB:GN00023; CESP:K04A8.5

C:Genetics:

A:Gene: K04A8.5

A:Map position: 5

C:Superfamily: triacylglycerol lipase, lingual

Query Match 30.8%; Score 692; DB 2; Length 411;

Best Local Similarity 36.8%; Pred. No. 7.5e-53;

Matches 153; Conservative 83; Mismatches 154; Indels 26; Gaps 10;

QY 11 VSHRNEMLLILVAYMFORNVNSVHMTKAVDPEAFNMNISEIIIOHGVPCPEYEVATEDG 70

DB 5 ISDLMTWPIILLISLNSY-----SKSVDFEYLDTPELIKSWGYSVEIYNTTKDG 57

QY 71 YILSVNRIPLGLVOPKKT---GSRPVVLLQHLGVGASNMISNLPNNSLGFILADAGDV 127

DB 58 FIELHRIPIYGREVPTSSDVNNRPVIFLQHGLCSFDWVNSPHQSAGVFADAGFDV 117

QY 128 WMGNSRGNASRKHKHTLSIDQ--EFWAFSYDEMAREFDLPAVINFLQKTQGEKIYVGVYS 186

DB 118 WLGFRGNTYSRKHVSLNPKDKPFWDSWDQISEYDLPAMIKALISGOESLYTGF 177

QY 187 QGSTMGTGFIAFSTMPDLAQKIKMYFALAPIATVKHAKSPGKTFLLPDMIMKG---F 240

DB 178 LGLTTFPAKLUSTDPKFSRKIKKYFALAPIGSIKHAHG---VFELGRHFGKDYBEYVKKH 234

QY 241 GKKEFLYQTRFLRQLVYILGQV--ILDQICSNIMLLGGFNMMNSRASVYAAHTLAG 299

DB 235 GSDELFGSSLLFKKIVKYTCGLFDTLEEFCSDTLLFIGHTANENWNOTRIPVYLAHTPAG 294

QY 300 TSQVNLHWSQVNSGELRAFDMGSETKNLEKCNQPTPVRYVRVDMT--VPTAMWTGGQDW 358

DB 295 SSSNVMAHLDMQFSYGVPTDFMGEE--KNLKAYGCKLPPQYNFTGCIADVPYLFWSDDW 353

QY 359 LSNPEDEV-KMLLSEVNLVYHKN--IPEWAHVDFIAGLDAPHRYMNEIHLMOQEE 411

DB 354 LSTKQDLLEETFLPAQLNSQVVGVSFRFRIENYHNLHFIMGTNAASQVYINVTGIILODD 409

RESULT 11

T20480

hypothetical protein F01G10.7 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000

C:Accession: T20480

R:Hembry, C.

submitted to the EMBL Data Library, October 1996

A:Reference number: Z19280

A:Accession: T20480

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-426 &lt;WIL&gt;

A:Cross-references: EMBL:Z81055; PIDN:CAB02896.1; GSPDB:GN00022; CESP:F01G10.7

A:Experimental source: Clone F01G10

C:Genetics:

A:Gene: CESP:F01G10.7

A:Map position: 4

A:Introns: 29/3; 64/1; 186/3; 347/2

C:Superfamily: triacylglycerol lipase, lingual

Query Match 29.2%; Score 655.5; DB 2; Length 426;

Best Local Similarity 35.9%; Pred. No. 1.3e-49;

Matches 140; Conservative 85; Mismatches 144; Indels 21; Gaps 9;

QY 20 LILVAYMFORNVNSVHMTKAVDPEAFNMNISEIIIOHGVPCPEYEVATEDGVILSVNRI 79

DB 6 LFLILSFSINL-----SLAIDDECYMTVPEIGKHFGYSEVHLVRTTDEYILELHRIP 59

QY 80 -RGLVQPKTKGSRPVVLLQHLGVGASNMISNLPNNSLGFILADAGDVWNGNSRGNAS 138

DB 60 CKQNEKCDRSKRPRIVPMQHLGLADGFSWIPNLANQSAGFVFADAGFDIWANSRGTPAS 119

QY 139 RKHTKLSIDODEWAFSYDEMAREFDLPAVINFLQKTQGEKIYVGVSGQTMGTGFIAPST 198

DB 120 QKHIGYGPENQKFWNTWQQMSEFDLTASVDLVKTKQEFLLYGLHSGQTMWIMFSRLAE 179

QY 199 MPDLAQKIKMYFALAPIATVKHAKSP---GTFKLLPDMIMKGILFGKKEP--LYQTRFL 252

DB 180 NPEFAKKIRHFHALAPVATVSHIGLFGLEFGTKELTYAEI---LLGLRPLSPISIPRTV 235

QY 253 RQLVYLCGVILDOICSNIMLLGGFNMMNSRASVYAAHTLAGTSVQVNLHWSQAV 312

DB 236 QKMISYMCSPFFQNICTLDIGFDG--NEKMFNSRSGVYLCHTPATSVKDLQHNIOLV 294

QY 313 NSGELRAFDMGSETKNLEKCNQPTPVRYVRVDMTVPVTAMTGGDWLSNPEDVK--MLLSE 371

DB 295 KSQTVSNFQDGTG--GNIIEYQPTPEYDVTQINTPTVLYWSRDDIADTQDIRDSILSK 353

QY 372 VTNLIYHKN--NIPENAHVDFIAGLDAPHRYM 400

DB 354 MNKTIAGSLPELPHYSMDHDFVFGTHAAFDLY 383

RESULT 12

BT0949

egg-specific protein - silkworm

C:Species: Bombyx mori (silkworm)

C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 26-Feb-1998

C:Accession: JT0949; A28527

R:Sato, Y.; Yamashita, O.

submitted to JIPID, September 1991

A:Reference number: JT0949

A:Accession: JT0949

A:Molecule type: DNA

A:Residues: 1-559 <SAT>

A:Experimental source: larva

A:Note: this protein is a homotrimer

R:Indrasith, L.S.; Sasaki, T.; Yamashita, O.

J. Biol. Chem. 263, 1045-1051, 1988

A:Title: A unique protease responsible for selective degradation of a yolk protein in *Bt*

A:Reference number: A28527; MUID:88087166; PMID:3275655

A:Accession: A28527

A:Molecule type: protein

A:Residues: 19-40;133-144, 'T',146-152;229-248 <IND>

C:Genetics:

A:Map position: 19

C:Keywords: egg yolk; homotrimer

F:132-133/Cleavage site: Lys-Asn (egg 30.5K cysteine proteinase) #status experimental

F:228-229/Cleavage site: Arg-Asp (egg 30.5K cysteine proteinase) #status experimental

Query Match 25.6%; Score 574.5; DB 2; Length 559;

Best Local Similarity 33.2%; Pred. No. 2.5e-42;

Matches 132; Conservative 82; Mismatches 152; Indels 31; Gaps 9;

QY 32 NSVH--MPTKAYDPEAF-MNISIIHQGPCREYEVATEDGYILSVNRIPGLVQPKK 87

DB 171 NTFHDAISETQRENEDFHLNATELLKKHQYVPEHTVATDDGYHLTVLRIP-PTHQTRD 229

QY 88 TGSRPVLLHGLVGGASNMISNPNNSLGFILADAGFDVNMGNSRGNAKSKHTLSID 147

DB 230 DKKKPVALLHGLLSADLLMGCPKSLAYMLSDAGYDVLGNVGNKYSRSHVSKHPA 289

QY 148 QDEFWAFSYDEMARFDLPVINFILQKTGQEKIYVYGYSOQTGTMGTFIAFSTMPELAOKIK 207

DB 290 LNDWFKESENDEIALHDLPAITDHYLDISGQERLHYIGHSQATTFALMSQSPSYNEKIV 349

QY 208 MYFALAPIATVKHAKS-----PCTKFL-LLPDMWTKGLFGKKEFLYQTRFLQLVIYL 259

DB 350 SMHALSDIVYNNYVRSPLFRMIAPTSKFYQIHQVGHGAFEPCKHLIET-----F 400

QY 260 CGQVILQD-----ICSNIMLLGGFNFTNNMNSRASYAAHTLAGTSVQNTILHWSQAVNS 314

DB 401 GGAACREKLCGRHVCNINLYVISINYNQDADIVPVMAHLPGATSARVMKQYQONVAS 460

QY 315 GELRAFQNGSTKLEKNQPTPVRYRVDMVTYPTAMWTGGQDWLSNPEDVKYMLLSEVTN 374

DB 461 HDFRKNYNGAET-NMKVYGASEPPSYDLSKVSAPVNLVHSHDAWLAKHPKQVEKLQENLPN 519

QY 375 LIYHKNPENAH---VDIFWGLDAPHRMYNEIIHLMQ 408

DB 520 VKQSFEVPEQHQHFDLDFQFSKKAPDTYVQKLMDNQ 556

RESULT 13

T39540

triacylglyceride lipase-cholesterol esterase - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Mar-2000

C:Accession: T39540

R:Wood, V.; Rajadream, M.A.; Barrell, B.G.; Beck, A.; Reinhardt, R.

submitted to the EMBL Data Library, February 1998

A:Reference number: 221862

A:Accession: T39540

A:Status: preliminary; translated from GB/EMBL/DBDJB

A:Molecule type: DNA



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:44:54 ; Search time 14 Seconds  
(without alignments)  
1253.178 Million cell updates/sec

Title: US-09-333-159-47

Perfect score: 2247

Sequence: 1 MLETLRWIVSHRMWMLL.....IHLMQQETNLISQGRCAVL 423

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1289	57.4	399	1 LICH_HUMAN	P38571 homo sapien
2	1220.5	54.3	397	1 LICH_RAT	Q64194 rattus norv
3	1197.5	53.3	397	1 LICH_MOUSE	Q920m5 mus musculu
4	1164.5	51.8	395	1 LIPG_RAT	P04634 rattus norv
5	1154	51.4	398	1 LIPG_CANFA	P80035 canis famil
6	1139	50.7	397	1 LIPG_BOVIN	Q29458 bos taurus
7	1139	49.8	398	1 LIPG_HUMAN	P07098 homo sapien
8	663	29.5	394	1 LIP3_DROME	O46108 drosophila
9	546	24.3	439	1 LIP1_DROME	O46107 drosophila
10	422.5	18.8	548	1 TGL1_YEAST	P34163 saccharomyc
11	126	5.6	987	1 YD94_METJA	Q58789 methanococc
12	105	4.7	430	1 ORC2_CAEEL	Q21037 caenorhabdi
13	101	4.5	436	1 PAPA_CAVPO	P70683 c platelet-
14	96.5	4.3	287	1 Y193_HAEIN	Q57427 haemophilus
15	95.5	4.3	1259	1 LIN1_HUMAN	P08547 homo sapien
16	93.5	4.2	278	1 PRXC_STRAU	O31168 streptomyc
17	93	4.1	569	1 PYND_PLAFA	Q08210 plasmodium
18	92	4.1	748	1 PTIP_ECOLI	P37177 escherichia
19	92	4.1	748	1 PTIP_SALTY	P37178 salmonella
20	91	4.0	370	1 P2C2_SCHPO	Q09172 schizosacch
21	90.5	4.0	1037	1 ACRD_ECOLI	P24175 escherichia
22	90	4.0	277	1 BPA2_STRAU	P29715 streptomyc
23	90	4.0	1122	1 YG3C_YEAST	P53280 saccharomyc
24	90	4.0	1136	1 POLG_OMV	P20234 o genome po
25	90	4.0	2083	1 DYSF_MOUSE	Q9esd7 mus musculu
26	89	4.0	330	1 EXOA_RHIME	P33691 rhizobium m
27	88.5	3.9	185	1 YCZH_BACSU	O31482 bacillus su
28	88.5	3.9	869	1 CFAC_ECOLI	P23733 escherichia
29	87	3.9	473	1 SCRB_LACLA	Q04937 lactococcus
30	86.5	3.8	382	1 YDHF_HAEIN	P44861 haemophilus
31	86.5	3.8	685	1 INVA_SALTY	P35657 salmonella
32	86.5	3.8	1882	1 Y468_MYCPN	P75109 mycoplasma
33	86	3.8	525	1 ACU8_NEUCR	P15937 neurospora

34	86	3.8	1866	1 VGNB_CPMV	P03500 cowpea mosa
35	85.5	3.8	428	1 YE63_SCHPO	O14249 schizosacch
36	85	3.8	429	1 ARE2_ECOLI	P52146 escherichia
37	85	3.8	837	1 UBPA_DICDI	P54201 dictyosteli
38	85	3.8	4128	1 PRKD_MOUSE	P97313 mus musculu
39	84.5	3.8	875	1 AMD2_XENLA	P12890 xenopus lae
40	84	3.7	451	1 YB27_YEAST	P38295 saccharomyc
41	84	3.7	804	1 SYL_BACSU	P36430 bacillus su
42	83.5	3.7	350	1 EL3B_PRUPE	P52408 prunus pers
43	83.5	3.7	375	1 NDR3_MOUSE	Q9qvf9 mus musculu
44	83.5	3.7	389	1 AAT_PYRAB	Q9v012 pyrococcus
45	83.5	3.7	456	1 YP95_YEAST	Q02891 saccharomyc

#### ALIGNMENTS

RESULT 1  
LICH\_HUMAN  
ID LICH\_HUMAN STANDARD: PRT: 399 AA.  
AC P38571; Q16529; Q96EJ0;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Lysosomal acid lipase/cholesteryl ester hydrolase precursor  
DE (EC 3.1.1.13) (LAL) (Acid cholesteryl ester hydrolase) (Sterol  
DE esterase) (Lipase A) (Cholesteryl esterase).  
GN LIPA.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP MEDLINE: FROM N.A., AND SEQUENCE OF 196-212; 277-297 AND 305-315.  
RX MEDLINE: 92042192; PubMed-1718995;  
RA Anderson R.A., Sando G.N.;  
RT "Cloning and expression of cDNA encoding human lysosomal acid  
RT lipase/cholesteryl ester hydrolase. Similarities to gastric and  
RT lingual lipases.";  
RL J. Biol. Chem. 266:22479-22484(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE: 94155897; PubMed-8112342;  
RA Ameis D., Merkel M., Eckerskorn C., Greten H.;  
RT "Purification, characterization and molecular cloning of human  
RT hepatic lysosomal acid lipase.";  
RL Eur. J. Biochem. 219:905-914(1994).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE: 96363957; PubMed-8725147;  
RA Du H., Witte D.P., Grabowski G.A.;  
RT "Tissue and cellular specific expression of murine lysosomal acid  
RT lipase mRNA and protein.";  
RL J. Lipid Res. 37:937-949(1996).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Placenta;  
RA Strausberg R.;  
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP VARIANT CSD/WD PRO-200.  
RX MEDLINE: 94195814; PubMed-8146180;  
RA Anderson R.A., Byrum R.S., Coates P.M., Sando G.N.;  
RT "Mutations at the lysosomal acid cholesteryl ester hydrolase gene  
RT locus in Wolman disease.";  
RL Proc. Natl. Acad. Sci. U.S.A. 91:2718-2722(1994).  
RN [6]  
RP VARIANT3 CSD ARG-129 AND PRO-129.  
RX MEDLINE: 98295576; PubMed-9633819;  
RA Ries S., Buechler C., Schindler G., Aslanidis C., Ameis D., Gasche C.,  
RA Jung N., Schambach A., Fehrer P., Vanier M.T., Belli D.C.,

RA Greten H., Schmitz G.;  
RT "Different missense mutations in histidine-108 of lysosomal acid  
RT lipase cause cholesteryl ester storage disease in unrelated compound  
RT heterozygous and hemizygous individuals.";  
RL Hum. Mutat. 12:44-51(1998).  
CC -1- FUNCTION: CRUCIAL FOR THE INTRACELLULAR HYDROLYSIS OF CHOLESTERYL  
CC ESTERS AND TRIGLYCERIDES THAT HAVE BEEN INTERNALIZED VIA RECEPTOR-  
CC MEDIATED ENDOCYTOSIS OF LIPOPROTEIN PARTICLES. IMPORTANT IN  
CC MEDIATING THE EFFECT OF LDL (LOW DENSITY LIPOPROTEIN) UPTAKE ON  
CC SUPPRESSION OF HYDROXYMETHYLGLUTARYL-COA REDUCTASE AND ACTIVATION  
CC OF ENDOGENOUS CELLULAR CHOLESTERYL ESTER FORMATION.  
CC -1- CATALYTIC ACTIVITY: A sterol ester + H(2)O -> a fatty  
CC acid.  
CC -1- SUBCELLULAR LOCATION: Lysosomal.  
CC -1- DISEASE: DEFECTS IN LIPA ARE THE CAUSE OF THE SEVERE INFANTILE-  
CC ONSET WOLMAN DISEASE (WD) AND THE Milder LATE-ONSET CHOLESTERYL  
CC ESTER STORAGE DISEASE (CESD).  
CC -1- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
CC EMBL; M74775; AAA59519.1; -  
CC EMBL; U04285; AAB60327.1; -  
CC EMBL; U04286; AAB60327.1; JOINED.  
CC EMBL; U04287; AAB60327.1; JOINED.  
CC EMBL; U04288; AAB60327.1; JOINED.  
CC EMBL; U04290; AAB60327.1; JOINED.  
CC EMBL; U04291; AAB60327.1; JOINED.  
CC EMBL; U04292; AAB60327.1; JOINED.  
CC EMBL; U04293; AAB60327.1; JOINED.  
CC EMBL; X76488; CAA54026.1; -  
CC EMBL; Z31690; CAA83495.1; -  
CC EMBL; U08464; AAB60328.1; -  
CC EMBL; BC012287; AAH12287.1; -  
CC PIR; A39315; A39315.  
CC Genew; HGNC:6611; LIPA.  
CC MIM; 278000; -  
CC InterPro; IPR000073; Abhydrolase.  
CC InterPro; IPR000734; Lipase.  
CC Pfam; PF00561; abhydrolase.1.  
CC PROSITE; PS00120; LIPASE\_SER; 1.  
KW Hydrolase; Lipid degradation; Glycoprotein; Signal; Lysosome;  
KW Disease mutation; Polymorphism.  
FT SIGNAL 1 21  
FT CHAIN 22 399  
FT  
FT  
FT ACT\_SITE 174 174  
FT ACT\_SITE 374 374  
FT CARBOHYD 36 36  
FT CARBOHYD 72 72  
FT CARBOHYD 101 101  
FT CARBOHYD 161 161  
FT CARBOHYD 273 273  
FT CARBOHYD 321 321  
FT VARIANT 16 16  
FT  
FT VARIANT 129 129  
FT  
FT VARIANT 129 129  
FT  
FT VARIANT 200 200  
FT  
FT CONFLICT 23 23  
FT CONFLICT 29 29  
FT SEQUENCE 399 AA; 45415 MW; 55F27391306B609A CRC64;

Query Match 57.48; Score 1289; DB 1; Length 399;  
Best Local Similarity 63.5%; Pred. No. 2.9e-104;  
Matches 235; Conservative 55; Mismatches 80; Indels 0; Gaps 0;  
QY 40 AVDPAPNISEIIHOHOGYPCPEYEVEATEDGYILSYNRIPIRGVLQPKTKGSRPVVLLQHG 99  
DB 28 AVDPETNNVSEIISYWGFPSEYLVETEDGYILCLNRIPIGRKNHSDKGPVVFQHG 87  
QY 100 LVGGASNMISNLPNNSLGLFILADAGFVMMGNSRGNWSRKHKTLSDIDQDEFWAFSYDEM 159  
DB 88 LLADSNVVTNLANSGLFILADAGFVMMGNSRGNWSRKHKTLSDVSDQDEFWAFSYDEM 147  
QY 160 ARFDLPVAVINILOKTQGEKIYVGYSGTGTWGTAFSTMPDELAKQIKMYFALAPIATVK 219  
DB 148 AKYDLPASINILNKTQGEQYVYVGHSGTGTGTFATFQIPELAKRIKMFALGPVASVA 207  
QY 220 HAKSPGTFFLLPDMWIKGLFGKKEFYOTRFLRQLVLYLGGVILDOICSNIMLLGGF 279  
DB 208 FCTSPMAKGLRPLDILKDLGDFKFLPQSAFLKWLGVTHVTHILKELCNLCFLLCGF 267  
QY 280 NTNNMNSRASVYAAHTLAGTSVQNLHWSQAVNSGELRAFDGSETKNLEKCNQPTVR 339  
DB 268 NERNLMSRVDVYITTHSPAGTSVQNLHWSQAVKFKQAFDWGSSAKNYHYNSYPT 327  
QY 340 YRVDMVTPTAMVTGGQDWLNPEDVKMLLSEVNLVYHKNIPENVAHVDFIWLGLDAPHRM 399  
DB 328 YNVKMLVPTAVWSGGHDLADVDVYNILLQTITNLVPHSTPEWEHLDFIWLGLDAPWRL 387  
QY 400 YNEIHLMQQ 409  
DB 388 YNKIINLMRK 397  
RESULT 2  
LICH\_RAT STANDARD; PRG; 397 AA.  
ID LICH\_RAT  
AC Q64194;  
DT 15-JUN-2002 (Rel. 41, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DE Lysosomal acid lipase/cholesteryl ester hydrolase precursor  
DE (EC 3.1.1.13) (LAL) (Acid cholesteryl ester hydrolase) (Sterol  
DE esterase) (Lipase A) (Cholesteryl esterase).  
GN LIPA OR LAL.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=96129534; PubMed=8576647;  
RA Nakagawa H., Matsubara S., Kuriyama M., Yoshidome H., Fujiyama J.,  
RT Yoshida H., Osame M.;  
RT "Cloning of rat lysosomal acid lipase cDNA and identification of the  
RT mutation in the rat model of Wolman's disease.";  
RL J. Lipid Res. 36:2212-2218(1995).  
CC -1- FUNCTION: CRUCIAL FOR THE INTRACELLULAR HYDROLYSIS OF CHOLESTERYL  
CC ESTERS AND TRIGLYCERIDES THAT HAVE BEEN INTERNALIZED VIA RECEPTOR-  
CC MEDIATED ENDOCYTOSIS OF LIPOPROTEIN PARTICLES. IMPORTANT IN  
CC MEDIATING THE EFFECT OF LDL (LOW DENSITY LIPOPROTEIN) UPTAKE ON  
CC SUPPRESSION OF HYDROXYMETHYLGLUTARYL-COA REDUCTASE AND ACTIVATION  
CC OF ENDOGENOUS CELLULAR CHOLESTERYL ESTER FORMATION (By  
CC similarity).  
CC -1- CATALYTIC ACTIVITY: A sterol ester + H(2)O -> a fatty  
CC acid.  
CC -1- SUBCELLULAR LOCATION: Lysosomal.  
CC -1- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way



```

Db 246 CTHVIMKELCANVFLCGLGKGNPNANMISQITVYPCQEVVTEGYILGV 60
QY 320 FDMGSEKLEKCNQPTPVRYVRDMVPTAMMTGGODWLSNPEDVKMLLSEVTNLVHK 379
Db 306 FDMGSSKNTFYHNSPPSYNIKNMLPTALWGGGRDLADINDITILLTQIPKLVYHK 365
QY 380 NIPWAHVDFIWLGLDAPHRMYNEIHLMOQ 409
Db 366 NIPEDWHLDFIWLGLDAPKLYDEIISLMKK 395

RESULT 4
LIPG_RAT
ID LIPG_RAT STANDARD; PRT; 395 AA.
AC P04634;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Triacylglycerol lipase, lingual precursor (EC 3.1.1.3) (Lingual
lipase).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85215587; PubMed=3839077;
RA Docherty A.J.P., Bodmer M.W., Angal S., Verger R., Riviere C.,
RA Lowe P.A., Lyons A., Entage J.S., Harris T.J.R.;
RT "Molecular cloning and nucleotide sequence of rat lingual lipase
cDNA."
RL Nucleic Acids Res. 13:1891-1903(1985).
CC -1- CATALYTIC ACTIVITY: Triacylglycerol + H(2)O -> diacylglycerol + a
fatty acid anion.
CC -1- MISCELLANEOUS: LINGUAL LIPASE IS SECRETED BY THE SEROUS (VON
EBNER'S) GLANDS AT THE BACK OF THE RAT TONGUE.
CC -1- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X02309; CAA26179.1; -
CC EMBL; A01157; CAA00136.1; -
CC PIR; A23045; LIRTT.
CC InterPro; IPR000073; Abhydrolase.
CC InterPro; IPR000734; Lipase.
CC InterPro; IPR000379; Ser_estrs_site.
CC Pfam; PF00561; abhydrolase.1.
CC PROSITE; PS00120; LIPASE_SER; 1.
KW Hydrolase; Lipid degradation; Glycoprotein; Signal.
FT SIGNAL 1
FT CHAIN 19 395
FT ACT_SITE 171 171
FT ACT_SITE 342 342
FT ACT_SITE 371 371
FT DISULFID 245 254
FT CARBOHYD 33 33
FT CARBOHYD 68 68
FT CARBOHYD 98 98
FT CARBOHYD 270 270
SQ SEQUENCE 395 AA; 44588 MW; E501854A923522EA CRC64;

Query Match 51.8%; Score 1164.5; DB 1; Length 395;
Best Local Similarity 54.98; Pred. No. 1.8e-93;
Matches 21; Conservative
QY 17 MWLLILVAYMFO-RNVNSVHMPTKAVDPAPMNISETIIHQYGPCEYEVEVATEDGYILSV 75

```

```

Db 1 MWLLITSVISTFGGAHGLGKLGPNANMISQITVYPCQEVVTEGYILGV 60
QY 76 NRIRGLVQPKTKTSRPPVLLQHLGVGASNWSNLNPNNSLPGFILDAGFDVWNGNSRGN 135
Db 61 YRIRPHGKNNSENIGKRPVYLQHLGIASATNWIANLPNNSLAFMLADAGYDVWILGNSRGN 120
QY 136 AWSRKHKTLSIDODEFWAFSYDEMARDLPVAVINFILQKTGOEKIYVYVYSGQTTMGFIA 195
Db 121 TWSRKNVYSPDSVEFWAFSDEMAKYDLPATINFIQVKTGOEKIHYVGHSGQTTMGFIA 180
QY 196 FSTMPDLAQIKMYFALAPIATVKHAKSPGTFKLLLPDMIKGLFGKKEFLYOTFRULQ 255
Db 181 FSTNPTLAKKIKTEFALAPVATVYKTSPLKKISFIPTEFLKFMGKMKMELPHTYFDDFL 240
QY 256 VIYLCGGVILDQICSNIMLLGGTNNMNSRASYVAAHTLAGTSVQNIHLWQAVNSG 315
Db 241 GTEVCSREVLDLCSNTLFIKGFDDKKNLNSRFDVYLGHNPAGTSVQDFLHWAQLVRS 300
QY 316 ELRAFDMGSEKLEKCNQPTPVRYVRDMVPTAMMTGGODWLSNPEDVKMLLSEVTNL 375
Db 301 KFOAFNMGSPQSNMLHYNQKTPPEYDVSAMTPVAVVWNGNDILADPDQDVAMLLPKLSNL 360
QY 376 IYHKNIPWAHVDFIWLGLDAPHRMYNEIHLMOQ 410
Db 361 LFHKEILAYNHLDFIWLADAPQEVYNEISMWAE 395

RESULT 5
LIPG_CANFA
ID LIPG_CANFA STANDARD; PRT; 398 AA.
AC P80035; O02857;
DT 01-NOV-1991 (Rel. 20, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Triacylglycerol lipase, gastric precursor (EC 3.1.1.3) (Gastric
lipase) (GL).
GN LIPF.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Stomach;
RA Vaganay S., Jolliff G., Bertaux O., Toselli E., Devignes M.D.,
RA Benicourt C.;
RA Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
(2)
SEQUENCE OF 20-59.
RX MEDLINE=92037652; PubMed=1935982;
RA Carriere F., Moreau H., Raphel V., Laugier R., Benicourt C.,
RA Junien J.-L., Verger R.;
RT "Purification and biochemical characterization of dog gastric
lipase."
RL Eur. J. Biochem. 202:75-83(1991).
CC -1- CATALYTIC ACTIVITY: Triacylglycerol + H(2)O -> diacylglycerol + a
fatty acid anion.
CC -1- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Y13899; CAA74198.1; -
CC PIR; S19539; S19539.
CC InterPro; IPR000073; Abhydrolase.
CC InterPro; IPR000734; Lipase.
CC InterPro; IPR000379; Ser_estrs_site.

```



DR Pfam: PF00561; abhydrolase; 1.  
DR PROSITE; PS00120; LIPASE\_SER; 1.  
KW Hydrolase; Lipid degradation; Glycoprotein; Signal.  
FT SIGNAL 1 19  
FT CHAIN 20 398 TRIACYLGLYCEROL LIPASE, GASTRIC.  
FT ACT\_SITE 172 172 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 372 372 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT CARBOHYD 34 34 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 99 99 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 327 327 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CONFLICT 39 39 I -> T (IN REF. 2).  
SQ SEQUENCE 398 AA; 45130 MW; E04D62F7518E386C CRC64;  
  
Query Match 51.4%; Score 1154; DB 1; Length 398;  
Best Local Similarity 54.3%; Pred. No. 1.5e-92;  
Matches 216; Conservative 71; Mismatches 99; Indels 12; Gaps 5;  
  
Qy 17 MWLLILVAYMFORNVNSVH-----MPTKAVDPFAFMNISEIIQHOGYPCEEYEVATEDG 70  
Db 1 MWLLTAAASVIS-TLGTTHGLFGKLHPT---NPEVTNNISQIMITYGWYPABEYEVATEDG 56  
  
Qy 71 YILSVNRIPRGLVQPKKTSRPRVLLQHLVGGASNWSLNPNNSLGFILADAGFDVWVG 130  
Db 57 YILGIDRIPYGRKNSENTGRPRVAFVLFQHLGLASATNWSLNPNNSLAFILADAGFDVWVG 116  
  
Qy 131 NSRGNAWRKHTLSIDODEWAFSYDEMAREFDLPVAVINFILOKTQGEKIYYVGSQGT 190  
Db 117 NSRGNTWARRNLYSPDSVEFWAFSDFEMAKYDLPATIDFILKTKQDQKLHVHGSQGT 176  
  
Qy 191 MGTAFTSMPELAQIKMYFALAPATVATVHKASPGTKFLLPDMMIKGLFCKEFLYQTR 250  
Db 177 IGTAFSTNPKLAKRIKTYFALAPVATVYKTYETTLNKLMLVPSFLFKIFGNKIP-YPHH 235  
  
Qy 251 FLRO-LVLYLGOVILDOICSNIMLLGGFNTNMNMSRASVYAAHTLAGTSVQNLHWS 309  
Db 236 FDDQLATEVCSRETVDLCSNALFIICGFDTMNLNMSRLDVLHNPAGTSVQNLHWS 295  
  
Qy 310 QAVNSGELRAFDMWGETKLEKCNQPTPVRYVRVDMTVPYTAAMTGGQDWSNPEDVKMLL 369  
Db 296 QAVKSGKFOAFDMGSPVQNMHHYQSMPPYNTLDMHVPYIAVWNGNDLLADPHDVL 355  
  
Qy 370 SEVTLNLIYKNIPEAHVDFIWLGLDAPHRMNEIHL 407  
Db 356 SKLPNLIIYHRKIPPYNHLDFIWMADPAQAVYNEIYSNM 393  
  
RESULT 6  
LIPG\_BOVIN STANDARD; PRT; 397 AA.  
ID LIPG\_BOVIN STANDARD; PRT; 397 AA.  
AC Q29458;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Triacylglycerol lipase, pregastric precursor (EC 3.1.1.3) (Pregastric  
lipase) (GL) (Pregastric esterase) (PGE).  
GN LIPP.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-Tongue serous gland;  
RX MEDLINE=95011625; PubMed=7926811;  
RA Timmermans M.Y.J., Kupers L.P., Teuchy H.;  
RT "The cDNA sequence encoding bovine pregastric esterase.";  
RL Gene 147:259-262(1994).  
CC -1- CATALYTIC ACTIVITY: Triacylglycerol + H(2)O -> diacylglycerol + a  
CC fatty acid anion.  
CC -1- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.  
-----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@sib-sib.ch).  
-----  
CC EMBL; L26319; AAA57037.1; -  
CC InterPro: IPR000073; Abhydrolase.  
CC InterPro: IPR000734; Lipase.  
CC InterPro: IPR000379; Ser\_estrs\_site.  
CC Pfam: PF00561; abhydrolase; 1.  
CC PROSITE; PS00120; LIPASE\_SER; 1.  
KW Hydrolase; Lipid degradation; Glycoprotein; Signal.  
FT SIGNAL 1 19  
FT CHAIN 20 397 TRIACYLGLYCEROL LIPASE, PREGASTRIC.  
FT ACT\_SITE 171 171 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 371 371 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT CARBOHYD 33 33 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 270 270 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 326 326 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 397 AA; 45231 MW; F68977DED585EE36 CRC64;  
  
Query Match 50.7%; Score 1139; DB 1; Length 397;  
Best Local Similarity 53.7%; Pred. No. 3e-91;  
Matches 212; Conservative 68; Mismatches 113; Indels 2; Gaps 2;  
  
Qy 18 WLLILVAYMFORNVNSVHPTKAVOPEAFMNISEIIQHOGYPCEEYEVATEDGYILSVNR 77  
Db 3 WLLVTVCFIHMSGNAFCFLGKIATKNPEASNMVSMYSYWGYPSEMHKVITADGYILOVYR 62  
  
Qy 78 IPRGLVQPKKTSRPRVLLQHLVGGASNWSLNPNNSLGFILADAGFDVWNGSNGNAW 137  
Db 63 IPRGKNNANHLGQRPVFLQHLGLSATNWSLNPNNSLGFILADAGFDVWNGSNGNTW 122  
  
Qy 138 SRKHTKLTSDODEWAFSYDEMAREFDLPVAVINFILOKTQGEKIYYVGSQGTMGFIAS 197  
Db 123 AQEHLIYSPDSPEFWAFSDFEMAEYDLPSTIDFILRTQKQLHYVHSGQTTIGFIAS 182  
  
Qy 198 TPELAQIKMYFALAPATVATVHKASPGTKFLLPDMMIKGLFCKEFLYQTRFQ-LV 256  
Db 183 TSPTLAELKIKVYALAPVATVYTKSLFNKLALPHLFKLIIFGDKMF-YPHTFLLSOFLG 241  
  
Qy 257 IYLCQVILDOICSNIMLLGGFNTNMNMSRASVYAAHTLAGTSVQNLHWSQAVNSGE 316  
Db 242 VEMCSRETLDVLCCKNALFAITGVNDKNFNNSRLDVIYAHNPAGTSVQNLHWSQAVNSGK 301  
  
Qy 317 LRAFDMWGETKLEKCNQPTPVRYVRVDMTVPYTAAMTGGQDWSNPEDVKMLLSEVTNLI 376  
Db 302 FOAFDMGAPYQNLHMYHQPTPIYNTAMNVPYAVNSADNLLADPDQDVFLLSKLSNLI 361  
  
Qy 377 YHKNIPWAHVDFIWLGLDAPHRMNEIHLMOQEE 411  
Db 362 YHKEIPYNHLDFIWMADPAQEVYNEIVSLMAEDK 396  
  
RESULT 7  
LIPG\_HUMAN STANDARD; PRT; 398 AA.  
ID LIPG\_HUMAN STANDARD; PRT; 398 AA.  
AC P07098;  
DT 01-APR-1988 (Rel. 07, Created)  
DT 01-APR-1988 (Rel. 07, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Triacylglycerol lipase, gastric precursor (EC 3.1.1.3) (Gastric  
lipase) (GL).  
GN LIPP.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.



```

RX MEDLINE=87299724; PubMed=3304425;
RA Bodmer M.W., Angal S., Yarranton G.T., Harris T.J.R., Lyons A.,
RA King D.J., Pieroni G., Riviere C., Verger R., Lowe P.A.;
RT "Molecular cloning of a human gastric lipase and expression of the
RT enzyme in yeast.";
RL Blochim. Biophys. Acta 909:237-244(1987).
RN [2]
RP SEQUENCE OF 20-45.
RX MEDLINE=89325292; PubMed=2753032;
RA Bernbaeck S., Blaackberg L.;
RT "Human gastric lipase. The N-terminal tetrapeptide is essential for
RT lipid binding and lipase activity.";
RL Eur. J. Biochem. 182:495-499(1989).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
RX MEDLINE=99287897; PubMed=10358049;
RA Roussel A., Canaan S., Egloff M.P., Riviere M., Dupuis L., Verger R.,
RA Cambillau C.;
RT "Crystal structure of human gastric lipase and model of lysosomal
RT acid lipase, two lipolytic enzymes of medical interest.";
RL J. Biol. Chem. 274:16995-17002(1999).
CC -1- CATALYTIC ACTIVITY: Triacylglycerol + H(2)O -> diacylglycerol + a
CC fatty acid anion.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE ABHDROLASE SUPERFAMILY. LIPASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X05997; CAA29413.1;
CC EMBL; X05997; CAA29414.1; ALT_INIT.
CC EMBL; A01046; CAA00125.1;
CC EMBL; A12714; CAA01053.1;
CC PIR; S04942; S04942.
CC PIR; S07145; S07145.
CC PDB; 1HLG; 15-MAR-00.
CC Genes; HGNC:6622; LIPF.
CC MTM; 601980;
CC InterPro; IPR000073; Abhydrolase.
CC InterPro; IPR000734; Lipase.
CC InterPro; IPR000379; Ser_estrs_site.
CC Pfam; PF00561; abhydrolase; 1.
CC PROSITE; PS00120; LIPASE_SER; 1.
CC Hydrolase; Lipid degradation; Glycoprotein; Signal; Polymorphism;
CC 3d-structure.
CC SIGNAL
CC CHAIN 1 19
CC ACT_SITE 20 398 TRIACYLGLYCEROL LIPASE, GASTRIC.
CC ACT_SITE 172 172 CHARGE RELAY SYSTEM.
CC ACT_SITE 343 343 CHARGE RELAY SYSTEM.
CC ACT_SITE 372 372 CHARGE RELAY SYSTEM.
CC DISULFID 246 255
CC CARBOHYD 34 34 N-LINKED (GLCNAC. . .).
CC CARBOHYD 99 99 N-LINKED (GLCNAC. . .).
CC CARBOHYD 271 271 N-LINKED (GLCNAC. . .).
CC CARBOHYD 327 327 N-LINKED (GLCNAC. . .).
CC VARIANT 161 161 T -> A (IN DBSNP:814628).
CC /FTID=VAR_011947.
CC SEQUENCE 398 AA; 45237 MW; CD3EE1621C014F0F CRC64;
Query Match 49.8%; Score 1119; DB 1; Length 398;
Best Local Similarity 52.2%; Pred. No. 1.6e-99;
Matches 210; Conservative 73; Mismatches 107; Indels 12; Gaps 2;
QY 17 MWLLILVAYFQFNVSVMHPTKRAV-----DPEAFMNISEIIHQGYPCEEYEVAED 69
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 MWLLLTWASLI-----SVLGTTHGLFCKLHGPSPEVTMNI SQMTYNGYPNEEYEVVTD 55
QY 70 GYLISVNRIPRGLVQPKTSRPSVLLQHLGVGGASWNISLNPNSLGLFIADAGFDVNM 129

```

```

Db 56 GYILEVNRIPYKKNKSGNTGQRPVFLQHGLLASATNWNISLNPNSLFIADAGYDVL 115
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 130 GNSRGNAWSKHKTLSDQDEFWAFSYDENARDLPVINFILQKTGOEKIYVYVSGQT 189
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 116 GNSRGNTWARRNLYYSPDSVEFWAFSEDEMAKYDLPATIDFIVKKTKQKOLHVGHSQGT 175
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 190 TMGFIASFMTPELAQIKMYFALAPATVYKHAQSPGTFLLLPDMIMKGLFGKKEFLYQT 249
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 176 TIGFIAFSTNPSLAKRIKTEYALAPVATVYKSLINKLRFVPSQLFKFEGDKIFVPHN 235
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 250 RFLROLVYILCGQVILDQICSNMLLLGGNTNNMNSRASVYAAHTLACTSVQNTLHWS 309
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 236 FFDQFLATEYCSREMLNLLCSNALFIICGDSKNFNTSLRDVYLSHNPACTSVQNMFWHT 295
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 310 QAVNSGELRAFDSGETKLEKNQPTPVRYRDMTPTAMMTGGODMLSNPDVKKML 369
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 296 QAVKSGKFAQYDWSGVPQNRHVDQSQPPYNTAMVPIAVWNGGKDLLADPDVGLLL 355
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 370 SEVTNLIYHKNPEMAHVDIFWGLDAPHRYNEIHLMOQEE 411
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 356 PKLPNLIYHKEIPFYHLDFTWMDAPQEVYNDIVSMISEDK 397
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
RESULT 8
LIP3_DROME
ID LIP3_DROME STANDARD; PRT; 394 AA.
AC O46108;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE "The Drosophila melanogaster lipase homologs: a gene family with
DE tissue and developmental specific expression."
DE Lipase 3 precursor (EC 3.1.1.-).
GN Lip3 OR CG8823.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Canton-S;
RX MEDLINE=98227315; PubMed=9566193;
RX Pistillo D., Manzi A., Tino A., Pilo Boyl P., Graziani F., Malva C.;
RT "The Drosophila melanogaster lipase homologs: a gene family with
RT tissue and developmental specific expression."
RN J. Mol. Biol. 276:877-885(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RX Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RX Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RX George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RX Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RX Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RX Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RX Abril J.F., Agrayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RX Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RX Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RX Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RX Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RX Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RX de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RX Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RX Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RX Fosler C., Gabriellista C.C., Garg N.S., Gelbart W.M., Glasser K.,
RX Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RX Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck C.,
RX Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RX Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RX Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RX Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RX Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

```

```
RA Merkulo G., Milshina N.V., Mobarrey C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
CC -1- TISSUE SPECIFICITY: FAT BODY.
CC -1- DEVELOPMENTAL STAGE: ONLY AT LARVAL STAGES.
CC -1- SIMILARITY: PARTIAL WITH OTHER LIPASES (PANCREATIC, GASTRIC,
CC HEPATIC, LINGUAL, LIPOPROTEIN, BACTERIAL, ETC.). ALSO SIMILAR TO
CC LEPIDOPTERAN EGG-SPECIFIC AND YOLK PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Y14367; CAA74737.1; ...
DR EMBL: AE003699; AAF54935.1; ...
DR FlyBase: FBgn0023495; Lip3.
DR InterPro: IPR000073; Abhydrolase.
DR InterPro: IPR000734; Lipase.
DR InterPro: IPR000379; Ser-estrase_site.
DR Pfam: PF00561; abhydrolase_1.
DR PROSITE: PS00120; LIPASE_SER; 1.
DR HydroLase: Lipid degradation; Signal: Glycoprotein.
FT SIGNAL 1 294
FT CHAIN 21 394
FT ACT_SITE 164 164
FT ACT_SITE 369 369
FT CARBOHYD 131 131
SQ SEQUENCE 394 AA; 44901 MW; A718D1D743673802 CRC64;
Query Match 29.5%; Score 663; DB 1; Length 394;
Best Local Similarity 37.0%; Pred. No. 5.8e-50;
Matches 137; Conservative 80; Mismatches 139; Indels 14; Gaps 7;
QY 51 EIIHQGYPCPEYEATGEGYILSVNIRPGLVQPK--KTGSRPVYLLQHLGLVGASNWI 108
DB 30 ERIEDGYPMERHEVVTSDNLTTHRIIP---YSPKTGESSNRPAFLMHGWLSSSDWV 86
QY 109 SNLPNNSLGLFIADAGFDVWNGNSRGNASHKHKLTSLDQDEFWAFSDYDEWAFDLPVAVI 168
DB 87 LMGPSRLAYMLADAGYDVWNGNARGNTYSKAHKWTPTYQIWFNFSNETHGMDVPAMI 146
QY 169 NFILQTKGQEKIYVGYSGQTMGFATSTPMLAQKIMVAFALAPIATVTKHAKSPGTK- 227
DB 147 DIVLAKTGOQVQVYGHSGQITVYLVYVSERPEYNDKIKSAHLLGPAAYMGKMSPLTRA 206
QY 228 ----FLLPDMTKLFGKKEFLYQYRFLRLQLVYLCGVY-LDQICSNIMLLGFGFTNN 283
DB 207 FAPILGQPNNAIVE-VCGSMEFPNSKFKODLGIEMCQATSPYADMCAEFLIGGYDTEQ 265
QY 284 MMSRASVYAATLACTSVQNTLHWSQVANGSELAFDWGSETKLEKCNQTPVRYRVR 343
DB 266 LDYELLEHIKATSPAGASVQNLHFCQYNSGKFRKFDY-TALRNPYEYGSYFPDYDLK 324
QY 344 DMTVPTAMWTGGDMLSPEDVKMLLSEVTNLIYHKNTIP--EWAHVDFIWLGLDAPRMYN 401
DB 325 NAKAPVLIYGGANDWCDVSDVRKRLDELPLNMAIDYLVFFKWAHLDFIWTGFEARKYVD 384
```

```
QY 402 EIIHLMQOEE 411
DB 385 EVLXQMOSYE 394
RESULT 9
LIP1_DROME STANDARD; PRT; 439 AA.
AC Q46107; Q9VRR6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lipase 1 precursor (EC 3.1.1.-).
GN Lip1 OR G7279.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE OF 7-433 FROM N.A.
RC STRAIN=Canton-S;
RX MEDLINE=98227315; PubMed=9566193;
RA Pistillo D., Manzi A., Tino A., Pilo Boyl P., Graziani F., Malva C.;
RT "The Drosophila melanogaster lipase homologs: a gene family with
RL tissue and developmental specific expression."
J. Mol. Biol. 276:877-885(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA DePalbis B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K.J., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,
RA Foaier C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel E.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mottel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulo G., Milshina N.V., Mobarrey C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
CC -1- FUNCTION: COULD BE A DIGESTIVE ENZYME.
CC -1- TISSUE SPECIFICITY: OVARIES.
CC -1- SIMILARITY: PARTIAL WITH OTHER LIPASES (PANCREATIC, GASTRIC,
CC HEPATIC, LINGUAL, LIPOPROTEIN, BACTERIAL, ETC.). ALSO SIMILAR TO
```

```
CC -----
CC LEPIDOPTERAN EGG-SPECIFIC AND YOLK PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Y14366; CAA74736.1; ..
CC EMBL; AE003629; AAF52994.1; ..
CC FlyBase; FBgn0023496; Lip1.
CC InterPro; IPR000073; Abhydrolase.
CC InterPro; IPR000734; Lipase.
CC InterPro; IPR000379; Ser_estrs_site.
CC Pfam; PF00561; abhydrolase; 1. FALSE_NEG.
CC PROSITE; PS00120; LIPASE_SER; Signal; Glycoprotein.
KW Hydrolyase; Lipid degradation; Signal; LIPASE 1.
FT SIGNAL 1 24
FT CHAIN 25 439
FT DOMAIN 30 44
FT ACT_SITE 197 197
FT ACT_SITE 393 393
FT CARBOHYD 124 124
FT CARBOHYD 151 151
FT CARBOHYD 346 346
FT CARBOHYD 379 379
FT CARBOHYD 426 426
FT CONFLICT 10 10
FT CONFLICT 213 213
FT CONFLICT 412 412
FT CONFLICT 439 439
SQ SEQUENCE 439 AA; 50660 MW; 9E32E20BEAE93E3F CRC64;

Query Match 24.38; Score 546; DB 1; Length 439;
Best Local Similarity 31.88; Pred. No. 9.6e-40;
Matches 123; Conservative 83; Mismatches 157; Indels 24; Gaps 7;

QY 39 KAVDPEAFMISEIHOQYPCPEYEVATEDGYILSVNRIPLGLVQPKTSRPPVILLOH 98
DB 58 KNIKQDSTLSVDKLIAYGYESEVHHVTTEDGYILTHRI-----RKQGA-PPFELLOH 109
QY 99 GLVGASNWSINLNNLSGLTADAGDVMGNSRGNASRKHKTLSIDQDEFWAFSYDE 158
DB 110 GLVDSSAGFVVMGPNVSLAYLLADHNYDVLGNARGNRSRHTTLDPEDESWHE 169
QY 159 MAREDLPAVINFILOKQGEKIYVYVSGQTMGFIAFSTMPDLAQIKMYFALAPIATV 218
DB 170 IGMVDLPAMDHVLKVTGFPKLHYAGHSQSGTSPFVCMSPAYNDKVVSMQALAPAVYA 229
QY 219 KHAKS-POTRFLLPDMKIGLFGCKEFLYOTRFLRLQILVYLCGOVILDOICSNIMLLIG 277
DB 230 KETEDHPYIRAIISLYNSLVGSSIRENFGEFLCRMT-----EETERLCIEAFVGI 283
QY 278 GFNTNNMNSRASYAAHTLAGTSVQNLHWSQAVNSGELRAFDGSGTKNLEKCNQPTP 337
DB 284 GRNNEFNRRKMFVILGHYPAGVAAKQVKHFIQIKSGRFAPYSY-SSNKNMQLYRDHLP 342
QY 338 VRYVRDNTVPTAMWTGGQDMLSPEDVKMLLSEVTNLIYKHNIP--EWAHVDFIWLDA 395
DB 343 PRYNLSLVPTFYVYSTNDLLCHPKDVESMCDLGNVTGKYLVPQKEFNHMDFLAIDV 402
QY 396 PHRYMNEIHLM-----QOQETNLSQ 416
DB 403 RKMLYRRMLQVLGVKPGCSPEANRSR 429

RESULT 10
TGL1_YEAST STANDARD; PRT; 548 AA.
AC P34163;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
```

```
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Triglyceride lipase-cholesterol esterase (EC 3.1.1.1.-).
GN TGL1 OR YKL140W OR YKL5.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92245761; PubMed=1574929;
RA Abraham P.R., Mulder A., Van T Riet J., Planta R.J., Raue H.A.;
RT "Molecular cloning and physical analysis of an 8.2 kb segment of
RT chromosome XI of Saccharomyces cerevisiae reveals five tightly linked
RT genes.";
RL Yeast 8:227-238 (1992).
RN [2]
RP SEQUENCE FROM N.A.
RA Rad M.R., Xu G., Kirchthath L., Fritz C., Keuchel H., Hollenberg C.P.;
RA Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
RL -1- SIMILARITY: BELONGS TO THE AB HYDROLASE SUPERFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z25464; CAA80958.1; ..
CC EMBL; Z28140; CAA81981.1; ..
CC PIR; S37969; S37969.
CC PIR; S39000; S39000.
CC SGD; S0001623; TGL1.
CC InterPro; IPR000073; Abhydrolase.
CC InterPro; IPR000379; Ser_estrs_site.
CC Pfam; PF00561; abhydrolase; 1.
KW Hydrolyase; Lipid degradation.
SQ SEQUENCE 548 AA; 62979 MW; 32D1F230701CB083 CRC64;

Query Match 18.88; Score 422.5; DB 1; Length 548;
Best Local Similarity 31.68; Pred. No. 6.7e-29;
Matches 119; Conservative 66; Mismatches 127; Indels 65; Gaps 15;

QY 61 EEEVATEDGYILSVNRIPLGLVOP--KKTGSRPVVLLQHLGVGASNWSINLP--NNSLG 117
DB 78 EDHLVREDNTYILTHRIIP-----PISKNRFNNKVVYLHGLMCSVDWCCNIEHKNLP 132
QY 118 FILADAGFDVWNGNSRGNASRKHKTLSIDQDEFWAFSYDEMARFDPVAVINFILQKTGQ 177
DB 133 FVLHDLGYDVMGNNRGNKYSTAHLNKKPKSNKFWDFSIDFAFFDIPNSTEFLDITKV 192
QY 178 EKIIYVYVSGQTTMGFTAFSTMPDLAQIKMYFALAPIATVK--HAK-----SPGTK 227
DB 193 DKVICIGFCGSAQMAFAFSLSEKLNKRVSHFIAIAPAMTPKGLHNRIVDPVLAKSSPGFM 252
QY 228 FLLLPDMKIGLFGCKEFLYQT----RFLRLQVILYLCGOVILDOICSNIMLLGENTNN 283
DB 253 YL-----FFGRKIVLPISAVIQWTLHTPTLFNLK-----ID-IANKIL-----FNWKS 293
QY 284 MNM-----SRASVYAAHTLAGTSVQNLHWSQAVNSGELRAFD-----WGSETKNLEK 331
DB 294 FNILPRQKIASYAKLYST-----TSVKSIVHWFQILRSQKFMQFEESDNMLNLSLRPYQI 348
QY 332 CNQPTPVRYVRDNTVPTAMWTGGQDMLSPEDVKMLLSEVTNLIYKHNIP--EWAHVDFIWL 391
DB 349 ANFPTRT-----NIKIPILYIGGIDSLVDIDVMMKNLP--FNSVFDVKVDNYEHLDLI 401
QY 392 GLDAPHRMYNEIHLMQ 408
DB 402 GKADATLVIAKVRFLFIE 418
```

```
RESULT 11
YD94_METJA          STANDARD;          PRT; 987 AA.
ID YD94_METJA
AC Q58789;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein MJ1394.
GN MJ1394.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Sult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutcliffe A.R., Dougherty B.A., Tomb J.F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT Jannaschii."
RL Science 273:1058-1073(1996).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: TO M.JANNASCHII MJ1393 AND A.FULGIDUS AF2028.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U67579; AAB99404.1;
CC TIGR: MJ1394;
CC KW Hypothetical protein; Transmembrane; Complete proteome.
CC FT TRANSMEM 12 32
CC FT TRANSMEM 958 978
CC FT SEQUENCE 987 AA; 112360 MW; D1E628FFB28CA86D CRC64;
CC
CC Query Match 5.6%; Score 126; DB 1; Length 987;
CC Best Local Similarity 23.1%; Pred. No. 0.0079;
CC Matches 105; Conservative 56; Mismatches 149; Indels 144; Gaps 29;
QY 3 ETLSRWIVSHRMEMWLLILVAYMFOR-----NVNSVHMPTKAVDPEAFMN-----48
DB 340 ETVDTVPKANKMSEVTNLVYLKLNLPNGENVNVNVIKYLNSGQFYDNFENWGWQ 399
QY 49 -----ISEIIHQHGYCEEVEVATEDGYILSVNRIPIRG--LVOPKKTGSRPVVLLQH 98
DB 400 YKNGIVQWSSIQSHSNYSLEKYGIST-----SLANDPNNGYKLLPKEIG-RDVI--S 450
QY 99 GLVGGASNWSINLPNNSLGFILADAGFDVWGMNSRGNASKHKT--LSIDQ-----148
DB 451 GWYRPSNW-GGGPIDRIG--LEDENF-----GYSEVNHYSNYSIDRTNGNPT 500
QY 149 ---DEFWAFSDYEMARFDLPVINFILOKTGQEKIYVYGSGCTMGTFATSTMPELAQ 205
DB 501 ISPEVYNPPDEWYFEL-----KI-----YSNCT-----ITFST-----531
QY 206 IKWYFALAPI-ATVKKIAKSGPTKF-----LLLPDMMIKG-----LFGKKEFLYQTR 250
DB 532 ---YQNGSLAATVSTIDNTYTFKDRVVIHGGYVYVDDLEVNSKNFDFYGDKNWYK--586
QY 251 FLRLVYILCGQVILQICSNMILLGG-----FNTNNMNSRASVYAAHTL-----AG 299
DB 587 -----EITSANSEGTAVLFDGDFYKDKDINTSLN---AINWNTITLWNSDSA 632
```

```
QY 300 TSVQNIL-HWSQAVNSGELRAFDWSETKNLEKCNQPTPVR--YVRVDMTVPVTAMWTGGQ 356
DB 633 TLVENVLGNYSYSERDNILAKYGFILFNNGTNTNTSIKGVYASGSYSISTDHGTGTE 692
QY 357 ---DWLSN---PEDVKMLLSEVNTLIYHKNIPEWA 385
DB 693 INIWTENVTFKNDAKSYSNFLTNL-----NI--WA 720
```

```
RESULT 12
ORC2_CAEEL
ID ORC2_CAEEL          STANDARD;          PRT; 430 AA.
AC Q21037;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Origin recognition complex subunit 2 (Ceorc2).
GN ORC-2 OR F59E10.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=96099401; PubMed=7502077;
RA Gavin K.A., Hidaka M., Stillman B.;
RT "Conserved initiator proteins in eukaryotes."
RL Science 270:1667-1671(1995).
RN [2]
SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Swinburne J.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
REVIEWS.
RA Jones S.J.M.;
RA Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: COMPONENT OF THE ORIGIN RECOGNITION COMPLEX (ORC) THAT
CC BINDS ORIGINS OF REPLICATION. IT HAS A ROLE IN BOTH CHROMOSOMAL
CC REPLICATION AND MATING TYPE TRANSCRIPTIONAL SILENCING. BINDS TO
CC THE ARS CONSENSUS SEQUENCE (ACS) OF ORIGINS OF REPLICATION IN AN
CC ATP-DEPENDENT MANNER.
CC -!- SUBUNIT: ORC IS COMPOSED OF SIX SUBUNITS (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: BELONGS TO THE ORC2 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U40270; AAC46954.1;
CC EMBL: Z36949; CAA85415.1;
CC WormPep: F59E10.1; CB11492.
CC DNA replication; Nuclear protein.
CC SEQUENCE 430 AA; 49319 MW; 54FC086BB4AD9670 CRC64;
CC
CC Query Match 4.7%; Score 105; DB 1; Length 430;
CC Best Local Similarity 21.9%; Pred. No. 0.17;
CC Matches 61; Conservative 42; Mismatches 87; Indels 88; Gaps 13;
QY 85 PKTGRPVVLLQHLGVGGASNWSINLPNN-----SLGF-----ILADAGFDW--MGN 131
DB 25 PEKESR-----QKTKNGKENASRNLSQNLDEDLQGLFEDETVTSQAIEYFQWQK 78
QY 132 S-----RGNWSRKHK-----TSLIDQDEFWAFSDYEMARFDLPVINFILOKTGQEKIY 181
DB 79 SASERNMNAKSRGRNRAGNCTEIEDEDEISNAITDFTKCDLPGLRNYTKKNDTEFEK 138
```

QY 182 YVGSQGTMGTAFTPEAQAQKMYFALAPATVYKHAQSPGKFLLLPDMKMGFLG 241  
 Db 139 RLEHLADNDFG-----KWKLYLAAG-----FNILLRGVGS 168  
 QY 242 KKEFLYQRFRLQVLYLGGVILDOICSNMILLGGFTN---NNMNSRASVYAAHTLA 298  
 Db 169 KRDVL--TEFENELSDYYMRVDARKDLNKKVLLGALNENKKNVKG-----217  
 QY 299 GTSVQNLHWSQAV-----NSGEL-----RAFDNGSE 325  
 Db 218 ---QSTISWARSIRKKNNSOOLILIDNIEAPDNRSD 251

RESULT 13  
 ID PAFA\_CAVPO STANDARD: PRT: 436 AA.  
 AC P70683;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Platelet-activating factor acetylhydrolase precursor (EC 3.1.1.47)  
 DE (PAF acetylhydrolase) (PAF 2-acylhydrolase) (LDL-associated  
 DE phospholipase A2) (LDL-PAF(2)) (2-acetyl-1-alkylglycerophosphocholine  
 DE esterase) (1-alkyl-2-acetyl-1-alkylglycerophosphocholine  
 GN PLAZG7 OR PAFAH.  
 OS Cavia porcellus (Guinea pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.  
 OX NCBI\_TaxID=10141;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Hartley; TISSUE=Liver;  
 RX MEDLINE=97103479; PubMed=8947850;  
 RA Karasawa K., Kuge O., Kawasaki K., Nishijima M., Nakano Y., Tomita M.,  
 RA Yokoyama K., Setaka M., Nojima S.;  
 RT Cloning, expression and characterization of plasma  
 RT platelet-activating factor-acetylhydrolase from guinea pig.";  
 RL J. Biochem. 120:838-844(1996).  
 CC -1- FUNCTION: MODULATES THE ACTION OF PLATELET-ACTIVATING FACTOR (PAF)  
 CC BY HYDROLYZING THE SN-2 ESTER BOND TO YIELD THE BIOLOGICALLY  
 CC INACTIVE LYSO-PAF. HAS A SPECIFICITY FOR SUBSTRATES WITH A SHORT  
 CC RESIDUE AT THE SN-2 POSITION. IT IS INACTIVE AGAINST LONG-CHAIN  
 CC PHOSPHOLIPIDS.  
 CC -1- CATALYTIC ACTIVITY: 2-acetyl-1-alkyl-sn-glycero-3-phosphocholine +  
 CC H(2)O = 1-alkyl-sn-glycero-3-phosphocholine + acetate.  
 CC -1- SUBCELLULAR LOCATION: Extracellular.  
 CC -1- TISSUE SPECIFICITY: PLASMA.  
 CC -1- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL: D67037; BAAL1054.1; -  
 DR InterPro: IPR000734; Lipase.  
 DR InterPro: IPR005065; PAF-AH\_P-II.  
 DR Pfam: PF03403; PAF-AH\_P-II; 1.  
 DR PROSITE: PS00120; LIPASE\_SER; 1.  
 KW Hydrolase; Lipid degradation; Glycoprotein; Signal.  
 FT SIGNAL 1 21  
 FT CHAIN 22 436  
 FT PLATELET-ACTIVATING FACTOR  
 FT ACETYLHYDROLASE.  
 FT ACT\_SITE 271 271  
 FT CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 294 294  
 FT CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 349 349  
 FT CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT CARBOHYD 76 76  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 200 200  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 324 324  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 436 AA; 49062 MW; C359D96E392FFE11 CRC64;  
 Query Match 4.5%; Score 101; DB 1; Length 436;  
 Best Local Similarity 20.4%; Pred. No. 0.38;  
 Matches 85; Conservative 60; Mismatches 149; Indels 122; Gaps 21;  
 QY 21 ILVAFMFORNVNSVMPKAVD-PEAFMNISEIIOHQGPCEEY-EVATED-----GY 71  
 Db 70 LMSGYTNQSSFLRLYYPSONDNDPDLW-----IPNEEYFQGLTETLGLASSFLGK 119  
 QY 72 ILSV---NRIPRGLVQPKKTSGR-PVYLLOHGLYGGASNMISNLPNNSGLFILADAGDF 126  
 Db 120 LLKLYGSKVPKAKWNSPLKTKGKYLIFSHGL--GAFRSI-----YSAIGIELASHGFI 173  
 QY 127 V-----WMNSRGNASWRKHKHTLSIDQDE-----FW 152  
 Db 174 VAAVEHRDESAAATYFYFADAPAEASGRNSRWIYKVCNLETERKRLQRGECQSALSW 233  
 QY 153 AFSYDEMARPDLPAVINFILOK-----TGQEKIYVYVYSGQTMGFIATFWPELAQKIM 208  
 Db 234 LLISIDEGEPVKNVLDNFDIQQLKGLSDRSKVAIIHGSFG--GATVIQTLSE--DQRFRC 289  
 QY 209 YFAL---APIATVKHAKSPGKFKLLPDMMIKGLFGKK-BELYOTRFRLRLV-----256  
 Db 290 GIALDPWMPVGEDVHSKIPOPLFFINSEYFOSANDTKKIEKFIQPKERKMAIVKGSVH 349  
 QY 257 -----IVLCGGOVLDOICSNIMILLGGFTNNMNN---SRASVYAAHTLACTSVQNLH 307  
 Db 350 HNFVDFTFAGKII-----GQMSLKGIKIDSEVADMLINKASLAFLOKYLGLD-KNFQ 402  
 QY 308 WSOAVNSGELRAFDWNGSETKLEKCNQPTPVRYRVRDVTPTAMTGGQDWLSNPE 363  
 Db 403 WNSLME-----GDDELN-----IPEFTIPTTMOSSGTGEORND 436

RESULT 14  
 ID Y193\_HAEIN STANDARD: PRT: 287 AA.  
 AC Q57427; O05013;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Putative esterase/lipase H10193 (EC 3.1.1.-).  
 GN H10193.  
 OS Haemophilus influenzae.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
 OC Haemophilus.  
 OX NCBI\_TaxID=727;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Rd / KW20 / ATCC 51907;  
 RX MEDLINE=95350630; PubMed=7542800;  
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
 RA Kervatage A.R., Sult C.J., Tomb J.F., Dougherty B.A., Merrick J.M.,  
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,  
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
 RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,  
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,  
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
 RA Venter J.C.;  
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae  
 Rd.";  
 RL Science 269:495-512(1995).  
 CC -1- SIMILARITY: BELONGS TO A FAMILY OF ESTERASES THAT GROUPS TOGETHER  
 CC PSEUDOMONA TROPINESTERASE, DMPD, TODF AND XYL.  
 CC -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

CC or send an email to license@isb-sib.ch).  
CC -----  
DR EMBL: U32704; AAC21862.1; -  
DR TIGR: H10193; -  
DR InterPro: IPR000073; Abhydrolase.  
DR InterPro: IPR000379; Ser\_estrif\_site.  
DR Pfam: PF00561; abhydrolase; 1.  
KW Hypothetical protein; Hydrolase; Serine esterase; Complete proteome.  
FT ACT\_SITE 119 119 BY SIMILARITY.  
FT ACT\_SITE 266 266 BY SIMILARITY.  
SQ SEQUENCE 287 AA; 32987 MW; F2B548619C060619 CRC64;

Query Match 4.3%; Score 95.5; DB 1; Length 287;  
Best Local Similarity 19.1%; Pred. NO. 0.53;  
Matches 62; Conservative 39; Mismatches 119; Indels 105; Gaps 12;

QY 84 QPKKTSRVRVLLHGLVGGASNSLNSPLADAGFDVWNGNSRGNWKRKHT 143  
DB 39 QVQKQINTVLIHGLFDNDN-----LGV-----ARAFSEHYI 75

QY 144 LSIQDDEFWAFSYDEMAREDPVAV- INFILQKTGOEKIYVYSGQTTMGFTAFSTMP 202  
DB 76 LRIDLNRHCHSPHSEKMYQLMAEDVIAVIRHLNLSKVLIGHSGMGTAMKITALCP 135

QY 203 AOKIKMYFALAPIATVVKHAKSPGTFKLLPDM-----MIKLFQKKEFLYQTR 250  
DB 136 VEK-----LIVIDHSPMPYEGFGHKDFVGLFAVYNAKPEN- 171

QY 251 FLROLVIYLCGOVILDOICSNIMLLGGFTNNMNSRASVVAHAHTLACTSVONILHWSQ 310  
DB 172 --RQAKPKLKEINDE--DVQFQMLKSFVNSADCFENL-----TALFNNTANMDWE- 222

QY 311 AVNSGELRFDWGSQTKNLEKCNQPTPVRYVRDVTPTAMTGGODMLSNPDVKMLIS 370  
DB 223 -----KVRVF-----TPLFIKGNSSVIKIENSEKILE 251

QY 371 EVTNLIYH--KNIPWAHV---DFI 390  
DB 252 QPNATAFTINGSGHWVHAEKPDFV 276

RESULT 15  
LIN1\_HUMAN  
ID LIN1\_HUMAN STANDARD; PRT; 1259 AA.  
AC P08547;  
DT 01-AUG-1988 (Rel. 08, Created)  
DT 01-AUG-1988 (Rel. 08, Last sequence update)  
DT 01-AUG-1988 (Rel. 08, Last annotation update)  
DE LINE-1 reverse transcriptase homolog.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=86230917; PubMed=2423883;  
RA Hattori M., Kuhara S., Takenaka O., Sakaki Y.;  
RT "Li family of repetitive DNA sequences in primates may be derived  
RT from a sequence encoding a reverse transcriptase-related protein."  
RL Nature 321:625-628(1986).  
CC -/- MISCELLANEOUS: THIS SEQUENCE WAS CONSTRUCTED FROM AN ALIGNMENT OF  
CC PUBLISHED AND UNPUBLISHED SEQUENCES, DETERMINED IN VARIOUS  
CC LABORATORIES, BELONGING TO THE LINE-1 FAMILY.  
DR PIR: A25313; GNHUL1.  
DR InterPro: IPR005135; Exo\_endo\_phos.  
DR InterPro: IPR000300; IPFC.  
DR InterPro: IPR000477; RVtse.  
DR Pfam: PF00078; rvt; 1.  
DR Pfam: PF03372; Exo\_endo\_phos; 1.  
DR SMART: SM00128; IPFC; 1.  
KW RNA-directed DNA polymerase.  
SQ SEQUENCE 1259 AA; 147112 MW; F3BC18A3803919F5 CRC64;

Query Match 4.3%; Score 95.5; DB 1; Length 1259;  
Best Local Similarity 19.7%; Pred. NO. 4.8;  
Matches 90; Conservative 67; Mismatches 170; Indels 129; Gaps 23;

QY 28 QRMNNSVHMPTKAVDPEAFMNI---SEIIQHOGYCEEYE----- 64  
DB 442 QBEVSLARPTTSSEIEAIIINSLPNKKSPEGTAETFYQRYKEELVFPFLKLQFSIEKE 501

QY 65 -VATEDGVILSVNRIIPR-GLVQPKTKTSRPVVL-----LQHGLVGGASNMISNL-PNN 114  
DB 502 GILPNSFEYEAIIILIPKGRDRTTKENFRPISLMNIDAKILANKILANQIQHKKLIHHD 561

QY 115 SLGFTLADAGFDVWNGNSR-----GNWRSRKKHTLSIDODEFWAFSYDEM----- 159  
DB 562 QVGFIPAMOG---WFNIRKSIINIIQHINRTKDTNHIISDAEK-----AFDKIQQPFMLK 614

QY 160 -----AREDLPAVINFILQKTGOEKIYVYSGQTTMGFTAFSTMP 203  
DB 615 PLMLKLGIDGTYUKIIRAIYDKPTA-NIIL--NQ-KLEAPPLKGTGRQCPPLSLPLNIV 670

QY 204 QKIKMYFALAPIATVVKHAKSPGTFKLLPDMMIKGLFGKKEFLYQTRFLRQLVYLCGOV 263  
DB 671 LEV---LA--RAIROEKE-----IKGIQLGKEEVKLSLFADDMIVYLENPI 711

QY 264 ILDQICSNIMLLGGFN-----TNNMNSRASVVAHAHTLACTSVONILHWSOAVNSGELR 318  
DB 712 VSNQ---NLKLISNFSKVSQKINVKQSAFLYTNROTESQINSELPTTIAASKRIKL 768

QY 319 AFDWGSQTKNLEKCNQPTPVRYVRDVT-----VPTAMTGGODMLSNPDVKMLISEVT 373  
DB 769 GTQLTRDVKDLPEKNY-KPLLNEIKEDTNNKKNIPCS-----WVGRINIVKMAI--LP 818

QY 374 NLIYHKN-IP-----EWAHVDFIWLGDAPH 397  
DB 819 KVIYRFNAIPKLPMTFTTELEKTLKLFIVNOKRAH 854

Search completed: February 3, 2003, 13:49:16  
Job time : 17 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:47:24 ; Search time 35 Seconds  
(without alignments)  
2490.227 Million cell updates/sec

Title: US-09-333-159-47

Perfect score: 2247  
Sequence: 1 MLTSLRQWIVSRMHWLL.....IHLMQBETNLSQRCVAVL 423

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL21.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriaph.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1991	88.6	374	4	Q96LG2
2	1134.5	50.5	395	11	Q9D798
3	1134.5	50.5	395	11	Q9D6X0
4	1134.5	50.5	395	11	Q9D6Q6
5	1134.5	50.5	395	11	Q9D6P3
6	1133.5	50.4	395	11	Q9CP88
7	1133.5	50.4	395	11	Q9CP77
8	1131.5	50.4	395	11	Q9D619
9	1129.5	50.3	395	11	Q9D8T5
10	1128.5	50.2	395	11	Q9D767
11	1127.5	50.2	395	11	Q9D7C5
12	1126.5	50.1	395	11	Q9D766
13	1126.5	50.1	395	11	Q9D611
14	1125.5	50.1	395	11	Q9D6Q3
15	1124.5	50.0	395	11	Q9D760
16	1124.5	50.0	395	11	Q9D655

17	1122.5	50.0	395	11	Q9D796
18	1119.5	49.8	395	11	Q9D6N8
19	1116.5	49.7	395	11	Q9D6Q2
20	1115.5	49.6	395	11	Q9D6M9
21	755.5	33.6	684	5	Q95XV1
22	752.5	33.5	403	5	Q61866
23	751	33.4	411	5	Q20449
24	730	32.5	405	5	Q93789
25	716	31.9	404	5	Q16956
26	715	31.8	169	11	Q9D2L7
27	692	30.8	411	5	Q94252
28	668.5	29.8	351	5	Q95X33
29	655.5	29.2	426	5	Q17766
30	655.5	29.2	434	5	Q9VKT9
31	639	28.4	456	5	Q9VK55
32	636	28.3	398	5	Q9V796
33	625.5	27.8	616	5	Q77107
34	588.5	26.2	838	5	Q9V8K6
35	574.5	25.6	559	5	Q17219
36	567.5	25.3	457	5	Q9VKT2
37	567.5	25.3	457	5	Q8T3X7
38	559	24.9	399	5	Q9VPE9
39	535	23.8	416	5	Q9VQ05
40	529.5	23.6	406	5	Q9VKT7
41	527	23.5	504	5	Q94568
42	505	22.5	443	5	P78898
43	502	22.3	435	5	Q9VG46
44	489	21.8	444	5	Q95U37
45	488	21.7	355	5	Q9VKT0

#### ALIGNMENTS

##### RESULT 1

Q96LG2 ID Q96LG2 PRELIMINARY; PRT: 374 AA.  
AC Q96LG2;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE BA30415.1 (Novel lipase) (Fragment).  
GN BA30415.1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Bray-Allen S.;  
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AL353113; CAC78754.1; -  
DR InterPro: IPR000073; Abhydrolase.  
DR InterPro: IPR000734; Lipase.  
DR InterPro: IPR000379; Ser\_estrs\_site.  
DR Pfam: PF00561; abhydrolase\_1.  
DR PROSITE, PS00120; LIPASE\_SER; UNKNOWN\_1.  
FT NON\_TER  
SQ SEQUENCE 374 AA; 42371 MW; 5E7220A889437337 CRC64;

Query Match 88.6%; Score 1991; DB 4; Length 374;  
Best Local Similarity 100.0%; Pred. No. 9e-167;  
Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 SEIIOHQYPCPEEYEVATEDGYILSVNRIPLGLVQPKKTSRPVLLQHLGLVGGASNWIS 109  
Db 1 SEIIOHQYPCPEEYEVATEDGYILSVNRIPLGLVQPKKTSRPVLLQHLGLVGGASNWIS 60  
Qy 110 NLPNNSLGFIADAGFDVWGMGNSRGNVNSRKHKTLSIDODEFWAFSYDEMARFDLPVAVN 169  
Db 61 NLPNNSLGFIADAGFDVWGMGNSRGNVNSRKHKTLSIDODEFWAFSYDEMARFDLPVAVN 120  
Qy 170 FLEQKTGQEKIYYVSGTMTMGFIATFMPLEAOKIKMYFALPATIATYKHAKSPTGKFL 229

Q9D796 mus musculus  
Q9D6N8 mus musculus  
Q9D6Q2 mus musculus  
Q9D6M9 mus musculus  
Q95XV1 caenorhabdi  
Q61866 caenorhabdi  
Q20449 caenorhabdi  
Q93789 caenorhabdi  
Q16956 caenorhabdi  
Q9D2L7 mus musculus  
Q94252 caenorhabdi  
Q95X33 caenorhabdi  
Q17766 caenorhabdi  
Q9VKT9 drosophila  
Q9VK55 drosophila  
Q9V796 drosophila  
Q77107 drosophila  
Q9V8K6 drosophila  
Q17219 bombyx mori  
Q9VKT2 drosophila  
Q8T3X7 drosophila  
Q9VPE9 drosophila  
Q9VQ05 drosophila  
Q9VKT7 drosophila  
Q94568 galliera me  
P78898 schizosacch  
Q9VG46 drosophila  
Q95U37 drosophila  
Q9VKT0 drosophila



```

Db 121 FILQKQGEKIYYVYSGQTTMGFIATFSTPELAQKIKMYFALAPIATVHKHAKSPGTRFL 180
Qy 230 LLPDMWIKGLFGKKEFLYQTRFLQOLYVYLCGOVILDOICSNIMLLGENTNNMNSRA 289
Db 181 LLPDMWIKGLFGKKEFLYQTRFLQOLYVYLCGOVILDOICSNIMLLGENTNNMNSRA 240
Qy 290 SVYAHAHTLAGTSVONILHWSQAVNSGELRAFDWGSSEKTKLEKCNQPTPVRYVRDVTPT 349
Db 241 SVYAHAHTLAGTSVONILHWSQAVNSGELRAFDWGSSEKTKLEKCNQPTPVRYVRDVTPT 300
Qy 350 AMWTGGQDNLSPEDVKMLLSEVNTLYHKNIPEVAHVDFIWLGLDAPHRMYNEIHLMOQ 409
Db 301 AMWTGGQDNLSPEDVKMLLSEVNTLYHKNIPEVAHVDFIWLGLDAPHRMYNEIHLMOQ 360
Qy 410 EETNLSQGRCAVL 423
Db 361 EETNLSQGRCAVL 374

RESULT 2
Q9D798 PRELIMINARY; PRT; 395 AA.
AC Q9D798;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE 2310051B21RIK protein.
GN 2310051B21RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TONGUE;
RX MEDLINE=21083660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaado I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staublei F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK009431; BAB26283.1;
DR MGD; MGI:1914967; 2310051B21RIK.
DR InterPro; IPR000073; Abhydrolase.
DR InterPro; IPR000734; Lipase.
DR Pfam; PF00561; abhydrolase; 1.
DR PROSITE; PS00120; LIPASE_SER; 1.
SQ SEQUENCE 395 AA; 44603 MW; D3FDB96FEA671E3E CRC64;

Query Match 50.5%; Score 1134.5; DB 11; Length 395;
Best Local Similarity 53.5%; Pred. No. 2.2e-91;
Matches 213; Conservative 70; Mismatches 108; Indels 7; Gaps 3;

Qy 17 MWLLTLVAYMFQ-RNVNSVHPTKAVDPEAFNINSEITQHQGPCEEVEATEDGYILSV 75
Db 1 MWLLLVTVLFAFGAGHGLFGKLPKNPEANMNVSQMITYGYPSEYEVVETDGYILGV 60

```

Query Match 50.5%; Score 1134.5; DB 11; Length 395;



```

Best Local Similarity 53.5%; Pred. No. 2.2e-91;
Matches 213; Conservative 70; Mismatches 108; Indels 7; Gaps 3;

Qy 17 MWLLILVAYFPQ-RNVNSVHMPTRKAVDPEAFMNISEIIHQGYPCPEEYEVATEDGYILSV 75
Db 1 MWLLLVTSVLSAFGAGHGLFGKLGPKNPEANMNVSMITYWGYPSEEEVYEDGYILGV 60

Qy 76 NRIPRGLVOPKTKSRPVLVQLHGLVGASNWSNLNPNLSLGLFIADAGFDVWNGSRGN 135
Db 61 YRIPYKGNSENICKRPVAYLQHLVASATNWTNLPNLSLAFILADAGYDVWNGSRGN 120

Qy 136 AWSKHKHTLSLTDQDEFWAFSDENARFDLPVINFILQKTOEKIYVYSGTGMGFA 195
Db 121 TWSRKNVYSPDSVEFWAFSDENAKYDLPATIDFIVQKTQEKIHYVGHSGQTGMGFA 180

Qy 196 FSTMPLEAKTKMYFALAPATVATVHKASPGTKFLLPDMIMIKGLFGKKEFL---YOTRFL 252
Db 181 FSTNPALAKKIKRYALAPVATVYTESPFKKISLIPKFLKLVIFGNKMFPHNYLDQFL 240

Qy 253 RQLVYLCGGVILDOICSNIMLLGGFNTNNMNSRASVYAAHTLAGTSVQNIHLWSQAV 312
Db 241 GTEV---CSRELLDLCSNALFICGFDKKNLNSRFDVYLGHNPAGTSTQDLFHWQA 297

Qy 313 NSGELRAFDMGSETKNLEKCNQPTPVRYVRDMVPTAMTGGQDMLSNPEDVKMLLSEV 372
Db 298 KSGKLOAYNWGSPQLQNLHYNQKTPPYDVSAMTVPIAVNNGGHDILADPODVAMLPLKL 357

Qy 373 TNLVYHKNIPEMAHVDFTWGLDAPHRMYNEIHLMQOE 410
Db 358 PNLLYHKELPYNHLDFIAMDAPQEVYNEIVTMMAED 395

RESULT 4
Q9D606 PRELIMINARY; PRT; 395 AA.
AC Q9D606;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Adult male tongue cDNA, RIKEN full-length enriched library,
DE clone:2310067K20, full insert sequence.
GN 2310051B21R1K.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TONGUE;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikola I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wyszew-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK010093; BAB26697.1;
DR MGD; MGI:1914967; 2310051B21R1K.
DR InterPro; IPR000073; Abhydrolase.

```

```

DR InterPro; IPR000734; Lipase.
DR InterPro; IPR000379; Ser_estrif_site.
DR Pfam; PF00561; abhydrolase_1.
DR PROSITE; PS00120; LIPASE_SER; 1.
SQ SEQUENCE 395 AA; 44623 MW; 245055F5E7FF0C9 CRC64;

Query Match 50.5%; Score 1134.5; DB 11; Length 395;
Best Local Similarity 53.8%; Pred. No. 2.2e-91;
Matches 214; Conservative 69; Mismatches 108; Indels 7; Gaps 3;

Qy 17 MWLLILVAYFPQ-RNVNSVHMPTRKAVDPEAFMNISEIIHQGYPCPEEYEVATEDGYILSV 75
Db 1 MWLLLVTSVLSAFGAGHGLFGKLGPKNPEANMNVSMITYWGYPSEEEVYEDGYILGV 60

Qy 76 NRIPRGLVOPKTKSRPVLVQLHGLVGASNWSNLNPNLSLGLFIADAGFDVWNGSRGN 135
Db 61 YRIPYKGNSENICKRPVAYLQHLVASATNWTNLPNLSLAFILADAGYDVWNGSRGN 120

Qy 136 AWSKHKHTLSLTDQDEFWAFSDENARFDLPVINFILQKTOEKIYVYSGTGMGFA 195
Db 121 TWSRKNVYSPDSVEFWAFSDENAKYDLPATIDFIVQKTQEKIHYVGHSGQTGMGFA 180

Qy 196 FSTMPLEAKTKMYFALAPATVATVHKASPGTKFLLPDMIMIKGLFGKKEFL---YOTRFL 252
Db 181 FSTNPALAKKIKRYALAPVATVYTESPFKKISLIPKFLKLVIFGNKMFPHNYLDQFL 240

Qy 253 RQLVYLCGGVILDOICSNIMLLGGFNTNNMNSRASVYAAHTLAGTSVQNIHLWSQAV 312
Db 241 GTEV---CSRELLDLCSNALFICGFDKKNLNSRFDVYLGHNPAGTSTQDLFHWQA 297

Qy 313 NSGELRAFDMGSETKNLEKCNQPTPVRYVRDMVPTAMTGGQDMLSNPEDVKMLLSEV 372
Db 298 KSGKLOAYNWGSPQLQNLHYNQKTPPYDVSAMTVPIAVNNGGHDILADPODVAMLPLKL 357

Qy 373 TNLVYHKNIPEMAHVDFTWGLDAPHRMYNEIHLMQOE 410
Db 358 PNLLYHKELPYNHLDFIAMDAPQEVYNEIVTMMAED 395

RESULT 5
Q9D6P3 PRELIMINARY; PRT; 395 AA.
AC Q9D6P3;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Adult male tongue cDNA, RIKEN full-length enriched library,
DE clone:2310069P19, full insert sequence.
GN 2310051B21R1K.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TONGUE;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikola I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wyszew-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK010093; BAB26697.1;
DR MGD; MGI:1914967; 2310051B21R1K.
DR InterPro; IPR000073; Abhydrolase.

```

RA Wyszaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
 RA Hayashizaki Y.,  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 DR EMBL; AK010139; BAB26725.1; -  
 DR MGD; MGI:1914967; 2310051B21Rik.  
 DR InterPro; IPR000073; Abhydrolase.  
 DR InterPro; IPR000734; Lipase.  
 DR Pfam; PF00561; abhydrolase.1.  
 DR PROSITE; PS00120; LIPASE\_SER; 1.  
 SQ SEQUENCE 395 AA; 44651 MW; E48A73CCCBFD359D CRC64;

Query Match 50.5%; Score 1134.5; DB 11; Length 395;  
 Best Local Similarity 53.8%; Pred. NO. 2.2e-91;  
 Matches 214; Conservative 69; Mismatches 108; Indels 7; Gaps 3;

QY 17 MWLLILVAYMFO-RNVNSVHMPTKAVDPFAFNISEIIOHQGYPCPEEYEVATEDGYLSV 75  
 DB 1 MWLLLVTSVLSAFGGAHGLFGKLGPKNPANMVNVSOMITYWGPSEEEYEVATEDGYLGV 60  
 QY 76 NRIPRGLVQPKTGRSPVYLLQHLVGGASNWSLNPNNSLGFILADAGFDVWNGNSRGN 135  
 DB 61 YRIPYKKNSENGRPAVYLOHGLIASATNITNLPNNSLAFILADAGYDVLGNSRGN 120  
 QY 136 AWSRKHKTLSDQDEFAFSDYEMAREFDLPVINFILQKTQGEKIYVYSGQTTMGFTA 195  
 DB 121 TWSRKNVYSPDSVEWAFSFDMAKYDLPATIDFIVQKTQGEKIHYVGHSGTTIGFTA 180  
 QY 196 FSTMPELAQIKMYFALAPIATVYKAKSPGTFFLLPDMWIKGLFGKKEFL---YQTRFL 252  
 DB 181 FSTNPALAKKIKRFYALAPIATVYKTESPFKISLPKFLKLVIFGNKMFPHNYLQDFL 240  
 QY 253 ROLVYLCGQVILDOICSNIMLLGGFNENNMNMSRASVYAAHTLAGTSVQNILLHWSQAV 312  
 DB 241 GTEV---CSRELLDLCSNALFICGDFKKNLNVSRFDVYLGNPAGTSTQDLFHWQA 297  
 QY 313 NSGELRAFDWGETNLEKCNQPTPVYRVYRDMVTPTAMTGGQWLSPEDVKMLLSEV 372  
 DB 298 KSGKLQAYNWGSPQLQNLHYNQKTPPYDVSANTVPIAVNGGHDILADPQDVAMLLPKL 357  
 QY 373 TNLVYHKNIPEWAHVDFTWGLDAPHRMYNEIHLMOOE 410  
 DB 358 PNLLYHKEILPYNHLDFIWADAPQEVYNEIVTMAED 395

RESULT 6  
 Q9CYP8 PRELIMINARY; PRT; 395 AA.  
 AC Q9CYP8;  
 DT 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE 2310051B21Rik protein.  
 GN 2310051B21Rik.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10990;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=TONGUE;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawaji J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Konno H., Adachi J., Fukuda S.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
 RA Wyszaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 DR EMBL; AK010116; BAB26711.1; -  
 DR EMBL; AK009300; BAB26201.1; -  
 DR MGD; MGI:1914957; 2310051B21Rik.  
 DR InterPro; IPR000073; Abhydrolase.  
 DR InterPro; IPR000734; Lipase.  
 DR Pfam; PF00561; abhydrolase.1.  
 DR PROSITE; PS00120; LIPASE\_SER; 1.  
 SQ SEQUENCE 395 AA; 44579 MW; D22996865EA671D34 CRC64;

Query Match 50.4%; Score 1133.5; DB 11; Length 395;  
 Best Local Similarity 53.5%; Pred. NO. 2.7e-91;  
 Matches 213; Conservative 70; Mismatches 108; Indels 7; Gaps 3;

QY 17 MWLLILVAYMFO-RNVNSVHMPTKAVDPFAFNISEIIOHQGYPCPEEYEVATEDGYLSV 75  
 DB 1 MWLLLVTSVLSAFGGAHGLFGKLGPKNPANMVNVSOMITYWGPSEEEYEVATEDGYLGV 60  
 QY 76 NRIPRGLVQPKTGRSPVYLLQHLVGGASNWSLNPNNSLGFILADAGFDVWNGNSRGN 135  
 DB 61 YRIPYKKNSENGRPAVYLOHGLIASATNITNLPNNSLAFILADAGYDVLGNSRGN 120  
 QY 136 AWSRKHKTLSDQDEFAFSDYEMAREFDLPVINFILQKTQGEKIYVYSGQTTMGFTA 195  
 DB 121 TWSRKNVYSPDSVEWAFSFDMAKYDLPATIDFIVQKTQGEKIHYVGHSGTTIGFTA 180  
 QY 196 FSTMPELAQIKMYFALAPIATVYKAKSPGTFFLLPDMWIKGLFGKKEFL---YQTRFL 252  
 DB 181 FSTNPALAKKIKRFYALAPIATVYKTESPFKISLPKFLKLVIFGNKMFPHNYLQDFL 240  
 QY 253 ROLVYLCGQVILDOICSNIMLLGGFNENNMNMSRASVYAAHTLAGTSVQNILLHWSQAV 312  
 DB 241 GTEV---CSRELLDLCSNALFICGDFKKNLNVSRFDVYLGNPAGTSTQDLFHWQA 297  
 QY 313 NSGELRAFDWGETNLEKCNQPTPVYRVYRDMVTPTAMTGGQWLSPEDVKMLLSEV 372  
 DB 298 KSGKLQAYNWGSPQLQNLHYNQKTPPYDVSANTVPIAVNGGHDILADPQDVAMLLPKL 357  
 QY 373 TNLVYHKNIPEWAHVDFTWGLDAPHRMYNEIHLMOOE 410  
 DB 358 PNLLYHKEILPYNHLDFIWADAPQEVYNEIVTMAED 395

RESULT 7  
 Q9CYP7 PRELIMINARY; PRT; 395 AA.  
 AC Q9CYP7;  
 DT 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE 2310051B21Rik protein.  
 GN 2310051B21Rik.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10990;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=TONGUE;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawaji J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Konno H., Adachi J., Fukuda S.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
RA Hayashizaki Y.,  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
DR EMBL: AK010236; BAB26787.1; -;  
DR EMBL: AK009413; BAB2672.1; -;  
DR EMBL: AK009428; BAB26280.1; -;  
DR EMBL: AK009459; BAB26300.1; -;  
DR EMBL: AK009473; BAB26312.1; -;  
DR EMBL: AK009474; BAB26313.1; -;  
DR EMBL: AK009479; BAB26316.1; -;  
DR EMBL: AK009523; BAB26338.1; -;  
DR EMBL: AK009525; BAB26339.1; -;  
DR EMBL: AK009546; BAB26352.1; -;  
DR EMBL: AK009571; BAB26368.1; -;  
DR EMBL: AK009573; BAB26370.1; -;  
DR EMBL: AK009729; BAB26466.1; -;  
DR EMBL: AK009773; BAB26495.1; -;  
DR EMBL: AK010019; BAB26647.1; -;  
DR EMBL: AK010035; BAB26656.1; -;  
DR EMBL: AK010058; BAB26673.1; -;  
DR EMBL: AK010061; BAB26675.1; -;  
DR EMBL: AK010124; BAB26715.1; -;  
DR EMBL: AK010125; BAB26716.1; -;  
DR MGD: MGI:1914967; 2310051B21Rik.  
DR InterPro: IPR000073; Abhydrolase.  
DR InterPro: IPR000734; Lipase.  
DR Pfam: PF00561; abhydrolase; 1.  
DR PROSITE: PS00120; LIPASE\_SER; 1.  
SQ SEQUENCE 395 AA; 44637 MW; D3F96B65EA671E34 CRC64;

Query Match 50.4%; Score 1131.5; DB 11; Length 395;  
Best Local Similarity 53.5%; Pred. No. 2.7e-91;  
Matches 213; Conservative 70; Mismatches 108; Indels 7; Gaps 3;

QY 17 MWLLILVAYMFQ-RNVNSVHMPKAVDPEAFPMNISEIIHQGYPCPEEYVATEDGYILSV 75  
Db 1 MWLLLVTSVLSAFSGAGHGLFGKLGPKNPEANNVSMITYGWYSEEVYEDGYILGV 60

QY 76 NRIPRGLVQPKKTSRPPVLLQHLGVGASNWSINLPNSLGLFLADAGFDVWNGSRGN 135  
Db 61 YRIPYKGNSENIGKRPVAVLQHLGLIASATNWTNLPNSLAFILADAGYDVLNLSRGN 120

QY 136 AWSRKHHTLSIDQDEFWAFSDENARFDLPVINFILQKTQGEKIYYGVSGQTMGFIA 195  
Db 121 TWSRKNVYSPDSVEFWAFSEDEMAKYDLPATIDFIVQKTQGEKIHYVGSQGTIGFIA 180

QY 196 FSTMPLEAOKIKMYFALAPATVTKHAKSPGTFKLLPDMIMKGLFGKKEFL---YQTRFL 252  
Db 181 FSTNPALAKKIKRYPALAPATVTKYTESPKKISLIPKELLKYIFGNKMFPHNYLDQFL 240

QY 253 RQLVIYLCGGVILQDQICSNIMLLGGFTNNMNSRVSVAHTLAGTSVQNIHLHWSQAV 312  
Db 241 GTEY---CSRLLDLCNSALFIFCGFDKKNLVSRFDVILGHNPAGTSTQDLFHWQA 297

QY 313 NSGEIRAFDGCSETKNLEKNCQTPPVRYRDMTPTVAMTGGQDNLSEDEYKMLLSEV 372  
Db 298 KSGKLQAINWGSPLQNNLHYNQKTPPYIDVSAMTVPVAVWNGGHDILADQDVAMLLPKL 357

QY 373 TNIYHKNIPEWAVHDFIWLGDAPHRMYNEIHLMOOE 410  
Db 358 PNLVYHKKEILPYNHLDFIWMADPAQEVYNEIVTWMAD 395

RESULT 8  
Q9D6L9  
ID Q9D6L9 PRELIMINARY: PRT; 395 AA.  
AC Q9D6L9;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Adult male tongue cDNA, RIKEN full-length enriched library,  
DE clone:2310076113, full insert sequence.  
GN 2310051B21Rik.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=TONGUE;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamahaka I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
RA Hayashizaki Y.,  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
DR EMBL: AK010203; BAB26766.1; -;  
DR MGD: MGI:1914967; 2310051B21Rik.  
DR InterPro: IPR000073; Abhydrolase.  
DR InterPro: IPR000734; Lipase.  
DR InterPro: IPR000379; Ser\_estrs\_site.  
DR Pfam: PF00561; abhydrolase; 1.  
DR PROSITE: PS00120; LIPASE\_SER; 1.  
SQ SEQUENCE 395 AA; 44671 MW; B8936162510AA55C CRC64;

Query Match 50.4%; Score 1131.5; DB 11; Length 395;  
Best Local Similarity 53.5%; Pred. No. 4e-91;  
Matches 213; Conservative 69; Mismatches 109; Indels 7; Gaps 3;

QY 17 MWLLILVAYMFQ-RNVNSVHMPKAVDPEAFPMNISEIIHQGYPCPEEYVATEDGYILSV 75  
Db 1 MWLLLVTSVLSAFSGAGHGLFGKLGPKNPEANNVSMITYGWYSEEVYEDGYILGV 60

QY 76 NRIPRGLVQPKKTSRPPVLLQHLGVGASNWSINLPNSLGLFLADAGFDVWNGSRGN 135  
Db 61 YRIPYKGNSENIGKRPVAVLQHLGLIASATNWTNLPNSLAFILADAGYDVLNLSRGN 120

QY 136 AWSRKHHTLSIDQDEFWAFSDENARFDLPVINFILQKTQGEKIYYGVSGQTMGFIA 195  
Db 121 TWSRKNVYSPDSVEFWAFSEDEMAKYDLPATIDFIVQKTQGEKIHYVGSQGTIGFIA 180

QY 196 FSTMPLEAOKIKMYFALAPATVTKHAKSPGTFKLLPDMIMKGLFGKKEFL---YQTRFL 252  
Db 181 FSTNPALAKKIKRYPALAPATVTKYTESPKKISLIPKELLKYIFGNKMFPHNYLDQFL 240

QY 253 RQLVIYLCGGVILQDQICSNIMLLGGFTNNMNSRVSVAHTLAGTSVQNIHLHWSQAV 312

Db	121	TWSRKNVYSPDSVEFWAFSEDEMAKYDLPATIDFVQKTQBEKIHVGHSGQTGTGFTA	181
Qy	196	FSTPELQAQIKMYFALAPATIAVKHAKSPGCTKFLLPDMXIKGLFGKKEFL---	252
Db	181	ESTPALAKKIKREYALAPYATVKYTESPEFKLSLPKPKLLKVFIGNKMFPHNYLDQEL	240
Qy	253	RLVYICGGVILDQICSNIMLLLGFTNTNMNMSRASVYAAHTLAGTSVQNTLHWSQAV	312
Db	241	GTEV---CSRELLDLGCSNALFIFCGFDKKNLNSRFDVYLGHNPAGTSTQDLFHWQA	297
Qy	313	NSGELRAFDSGSETKLEKCNQPTPVYRYRDMTVPMTAKMTGGDLSNPEDVKMLLSEV	372
Db	298	KSGKLQAYNNGSPQNMLHYNQSPPYDVVSAMTVPYAVWNGGHDILAQQDVAMLLPKL	357
Qy	373	TNLHYKNINPEAHVDFIWLGDAPHRMYNEIHLMOQE	410
Db	358	PNLLYHKEILPYNHLDFIWMADAPQEVYNEIVTMAED	395
RESULT 10			
ID	Q9D767	PRELIMINARY;	PRT; 395 AA.
AC	Q9D767;		
DT	01-JUN-2001	(T-EMBLrel. 17, Created)	
DT	01-JUN-2001	(T-EMBLrel. 17, Last sequence update)	
DT	01-MAR-2002	(T-EMBLrel. 20, Last annotation update)	
DE	2310051B21RIK	protein.	
GN	2310051B21RIK		
OS	Mus musculus (House).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxId=10090;		
RP	SEQUENCE FROM H.A.		
RC	STRAIN=C57BL/6J; TISSUE=TONGUE;		
RX	MEDLINE=21085660; PubMed=11217851;		
RA	Kawai J., Shingawara A., Shibata K., Yoshino M., Itoh M., Ishii Y.,		
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,		
RA	Alzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,		
RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,		
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,		
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,		
RA	Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,		
RA	Schulz L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,		
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,		
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,		
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,		
RA	Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,		
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,		
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,		
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,		
RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,		
RA	Yunshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,		
RA	Hayashizaki Y.		
RT	"Functional annotation of a full-length mouse cDNA collection."		
RL	Nature 409:685-690(2001).		
DR	EMBL; AK009537; BAB26346.1; -		
DR	MGP; MGI:1914967; 2310051B21RIK.		
DR	InterPro; IPR000073; Abhydrolase.		
DR	InterPro; IPR000734; Lipase.		
DR	InterPro; IPR000379; Ser_estrs_site.		
DR	Pfam; PF00561; abhydrolase.1.		
DR	PROSITE; PS00120; LIPASE_SER; 1.		
SQ	SEQUENCE 395 AA; 44607 MW; B43317C2254FA8FB CRC64;		
Query Match	50.2%;	Score 1128.5;	DB 11; Length 395;
Best Local Similarity	53.3%;	Pred. No. 7.4e-91;	
Matches	212; Conservative	Mismatches 109;	Indels 7; Gaps
Qy	17: MWLLILVAYNFQ-RNVNSVHMPTKRAVDPEAFPMNTSEIIHQGYPCREYEVATEDGYILSV	75	
Db	1 MWLLLVSVLSAFGAGHGLFKGLGPKPNANNSQMTWTGVPSEETVEDGYILGV	60	

QY 76 NRIPRGLVOPKTKGSRPVVLLQHLGVGGASNNISNPNNSLGFILADAGFDVWNGSRGN 135  
 DB 61 YRIPYKKNSENIGKRPVAYLQHLGLASATNWTNPNNSLAFILADAGFDVWNGSRGN 120  
 QY 136 AWRKHKHTLSIDODEWAFSDYDEMARFDPVAVNFILQTKGQEKIYVYSGQTTMGFTA 195  
 DB 121 TWSRKNVYSPDSVEFWAFSDYDEMARFDPVAVNFILQTKGQEKIYVYSGQTTMGFTA 180  
 QY 196 FSTMPLEAKIKMYFALAPATATVKHAKSPGKTKFLLPDMMIKGLFGKKEFL---YOTRFL 252  
 DB 181 FSTNPALAKIKRYFALAPVATVKYTESPKKISLPKLLKVFIGNKMFPHNYLDQFL 240  
 QY 253 RQLVYLCGOVILQICSNIMLLGGFNTNMMNSRASVYAAHTLAGTSVQNILHWSQAV 312  
 DB 241 GTEV---CSRELLDLCCSNALFIFCGFDKKNLNSRFDVYLGHNPAAGTSTQDLFWAQLA 297  
 QY 313 NSGELRAFDFWGETKLEKCNQPTPVRYRYRDMTPTAMTGGQDWLSNPDVYKMLLSEV 372  
 DB 298 KSGKLAQYNGWSPLONLHYNQKTPPYDVVSAMTVPVAVWNGGHDILADQDVAMLLPKL 357  
 QY 373 TNLTYHKNIPEWAHVDFIWLGLDAPHRMYNEIHLMOOE 410  
 DB 358 PNLLYHKKEILPYNHLDFIWMADAPQEVYNEIVTMMAED 395

RESULT 11

Q9D7C5 PRELIMINARY; PRT; 395 AA.  
 AC Q9D7C5;  
 DT 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE 2310051B21RIK protein.  
 GN 2310051B21RIK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=TONGUE;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli I., Sakamoto N.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sakai K., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 DR EMBL: AK009359; BAB26240.1; -;  
 DR MGD: MGI:1914967; 2310051B21RIK.  
 DR InterPro: IPR000073; Abhydrolase.  
 DR InterPro: IPR000734; Lipase.  
 DR Pfam: PF00561; abhydrolase; 1.  
 DR PROSITE: PS00120; LIPASE\_SER; 1.  
 SQ SEQUENCE 395 AA; 47478 MW; 3ED43992458DE058 CRC64;

Query Match 50.2%; Score 1127.5; DB 11; Length 395;  
 . Best Local Similarity 53.3%; Pred. No. 9.1e-91;

Matches 212; Conservative 69; Mismatches 110; Indels 7; Gaps 3;  
 QY 17 MWLLILVAYMFO--RNYSVHMPTKAVDPEAFMNISETIOHOGYPCPEYEVATEDGYILSV 75  
 DB 1 MWLLIVTSVLSAFGGAGHGLGKLPKNPEANMNVSMQITWGPSEYEVATEDGYILSV 60  
 QY 76 NRIPRGLVOPKTKGSRPVVLLQHLGVGGASNNISNPNNSLGFILADAGFDVWNGSRGN 135  
 DB 61 YRIPYKKNSENIGKRPVAYLQHLGLASATNWTNPNNSLAFILADAGFDVWNGSRGN 120  
 QY 136 AWRKHKHTLSIDODEWAFSDYDEMARFDPVAVNFILQTKGQEKIYVYSGQTTMGFTA 195  
 DB 121 TWSRKNVYSPDSVEFWAFSDYDEMARFDPVAVNFILQTKGQEKIYVYSGQTTMGFTA 180  
 QY 196 FSTMPLEAKIKMYFALAPATATVKHAKSPGKTKFLLPDMMIKGLFGKKEFL---YOTRFL 252  
 DB 181 FSTNPALAKIKRYFALAPVATVKYTESPKKIFKLLKVFIGNKMFPHNYLDQFL 240  
 QY 253 RQLVYLCGOVILQICSNIMLLGGFNTNMMNSRASVYAAHTLAGTSVQNILHWSQAV 312  
 DB 241 GTEV---CSRELLDLCCSNALFIFCGFDKKNLNSRFDVYLGHNPAAGTSTQDLFWAQLA 297  
 QY 313 NSGELRAFDFWGETKLEKCNQPTPVRYRYRDMTPTAMTGGQDWLSNPDVYKMLLSEV 372  
 DB 298 KSGKLAQYNGWSPLONLHYNQKTPPYDVVSAMTVPVAVWNGGHDILADQDVAMLLPKL 357  
 QY 373 TNLTYHKNIPEWAHVDFIWLGLDAPHRMYNEIHLMOOE 410  
 DB 358 PNLLYHKKEILPYNHLDFIWMADAPQEVYNEIVTMMAED 395

RESULT 12

Q9D766 PRELIMINARY; PRT; 395 AA.  
 AC Q9D766;  
 DT 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE 2310051E21RIK protein.  
 GN 2310051E21RIK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=TONGUE;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli I., Sakamoto N.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sakai K., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 DR EMBL: AK009544; BAB26350.1; -;  
 DR MGD: MGI:1914967; 2310051B21RIK.  
 DR InterPro: IPR000073; Abhydrolase.  
 DR InterPro: IPR000734; Lipase.  
 DR InterPro: IPR000379; Ser\_estrs\_site.

DR Pfam; PF00561; abhydrolase; 1.  
DR PROSITE; PS00120; LIPASE\_SER; 1.  
SQ SEQUENCE 395 AA: 44648 MW; CC69875653AA7A7A CRC64;

Query Match 50.1%; Score 1126.5; DB 11; Length 395;  
Best Local Similarity 53.3%; Pred. No. 1.le-90;  
Matches 212; Conservative 70; Mismatches 109; Indels 7; Gaps 3;

QY 17 MWLLLVAYMFO-RNVNSVHMTKAVDPEAFWNISIIIOHQPCEYEYEVATEDGYLSV 75  
DB 1 MWLLLVTSVLSAFGAHGLFGKLPKNPANNNVSNOMITYGYPSEYEVATEDGYLGV 60  
QY 76 NRIPRGLVOPKTKGRPVVLLQHLGVGGASNNISLNPNNSLGFLADAGDFVMMGNSRG 135  
DB 61 YRIPYKKNSEINIGKRPVAYLQGLIASATNITNLPNNSLAFILADAGYDVLGNSRG 120  
QY 136 ANSRKHTLSIDODEFAESYDEMAREFDLPVINFILQKTGOEKIYVYSGTGTGFTA 195  
DB 121 TWSRKNVYSPDSVEFAESFDEMAKYDLPATIDFVQKTGOEKIHYVHSGTGTGFTA 180  
QY 196 FSTMPLEAQIKMYFALAPIATVTKHAKSPGTFLLLPDMKIFLGKKEFL---YOTREL 252  
DB 181 FSTNPALAKKIRFYALAPVATVKYTESPFKISLIPKELLVIFGNKMFMPNYLDQFL 240  
QY 253 ROLVYILCGQVILDOICSNIMLLGFGTNNMNSRASVYAAHTLAGTSVQNTLHWSQAV 312  
DB 241 GTEV---CSRELLDLCSNALFIFCGFDKKNLVSDFVYLGHNPAGTSTQDLFHWQA 297  
QY 313 NSGELRAFQWSETKNLEKNQPTPVRYVRDMTPTAMTGGQDWSLNPEDVKMLLSEV 372  
DB 298 KSGKLOAYNWGSPQLNHLHYNOKTPPYDVVSAMTVPVIAVNWNGHDLADPQDVAMLLPKL 357

QY 373 TNLVYHKNIPEWAHVDFIWLGDAPHRYNEIHLMOQE 410  
DB 358 PNLLYHKELPYNHLDFIWMADAPQEVYNEIVTMAED 395

RESULT 13  
Q9D6L1  
ID Q9D6L1 PRELIMINARY; PRT; 395 AA.  
AC Q9D6L1;  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DE Adult male tongue cDNA, RIKEN full-length enriched library,  
DE clone:2310079020, full insert sequence.  
GN 2310051B21RIK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=TONGUE;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Glasi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher N., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamaya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai J., Kohtsuki S.,  
RA Hayashizaki Y.;

RT \*Functional annotation of a full-length mouse cDNA collection.\*;  
RL Nature 409:685-690(2001).  
DR EMBL; AK010231; BAB26784.1; -.  
DR MGD; MGI:1914967; 2310051B21RIK.  
DR InterPro; IPR000073; Abhydrolase.  
DR InterPro; IPR000734; Lipase.  
DR Pfam; PF00561; abhydrolase; 1.  
DR PROSITE; PS00120; LIPASE\_SER; 1.  
SQ SEQUENCE 395 AA: 44659 MW; 13E0BB95FC5A40E0 CRC64;

Query Match 50.1%; Score 1126.5; DB 11; Length 395;  
Best Local Similarity 53.5%; Pred. No. 1.le-90;  
Matches 213; Conservative 69; Mismatches 109; Indels 7; Gaps 3;

QY 17 MWLLLVAYMFO-RNVNSVHMTKAVDPEAFWNISIIIOHQPCEYEYEVATEDGYLSV 75  
DB 1 MWLLLVTSVLSAFGAHGLFGKLPKNPANNNVSNOMITYGYPSEYEVATEDGYLGV 60  
QY 76 NRIPRGLVOPKTKGRPVVLLQHLGVGGASNNISLNPNNSLGFLADAGDFVMMGNSRG 135  
DB 61 YRIPYKKNSEINIGKRPVAYLQGLIASATNITNLPNNSLAFILADAGYDVLGNSRG 120  
QY 136 ANSRKHTLSIDODEFAESYDEMAREFDLPVINFILQKTGOEKIYVYSGTGTGFTA 195  
DB 121 TWSRKNVYSPDSVEFAESFDEMAKYDLPATIDFVQKTGOEKIHYVHSGTGTGFTA 180  
QY 196 FSTMPLEAQIKMYFALAPIATVTKHAKSPGTFLLLPDMKIFLGKKEFL---YOTREL 252  
DB 181 FSTNPALAKKIRFYALAPVATVKYTESPFKISLIPKELLVIFGNKMFMPNYLDQFL 240  
QY 253 ROLVYILCGQVILDOICSNIMLLGFGTNNMNSRASVYAAHTLAGTSVQNTLHWSQAV 312  
DB 241 GTEV---CSRELLDLCSNALFIFCGFDKKNLVSDFVYLGHNPAGTSTQDLFHWQA 297  
QY 313 NSGELRAFQWSETKNLEKNQPTPVRYVRDMTPTAMTGGQDWSLNPEDVKMLLSEV 372  
DB 298 KSGKLOAYNWGSPQLNHLHYNOKTPPYDVVSAMTVPVIAVNWNGHDLADPQDVAMLLPKL 357

QY 373 TNLVYHKNIPEWAHVDFIWLGDAPHRYNEIHLMOQE 410  
DB 358 PNLLYHKELPYNHLDFIWMADAPQEVYNEIVTMAED 395

RESULT 14  
Q9D6Q3  
ID Q9D6Q3 PRELIMINARY; PRT; 395 AA.  
AC Q9D6Q3;  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE Adult male tongue cDNA, RIKEN full-length enriched library,  
DE clone:2310068C02, full insert sequence.  
GN 2310051B21RIK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=TONGUE;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Glasi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher N., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamaya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai J., Kohtsuki S.,  
RA Hayashizaki Y.;



```
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayaehizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK010103; BAB26703.1; -
DR MGI; MGI:1914967; 2310051B21Rik.
DR InterPro; IPR000073; Abhydrolase.
DR InterPro; IPR000734; Lipase.
DR InterPro; IPR000379; Ser_estrs_site.
DR Pfam; PF00561; abhydrolase_1.
DR PROSITE; PS00120; LIPASE_SER; 1.
SQ SEQUENCE 395 AA; 44638 MW; 4E52613BC9B8B3F CRC64;

Query Match 50.18; Score 1125.5; DB 11; Length 395;
Best Local Similarity 53.3%; Pred. No. 1.4e-90;
Matches 212; Conservative 69; Mismatches 110; Indels 7; Gaps 3;

QY 17 MWLLILVAYMFO-RNVSVMHPTKAVDPPEAFMNISEIIIOHOGYPCPEEVEVATEDGYILSV 75
DB 1 MWLLIVTSVLSAFSGAGHGLFGKLPKPNPEANNVSMITYWGYPSEYEVVTEGYILGV 60
QY 76 NRIPRGLVQPKTKTSRPPVYLQHLGVGGASNMISLNPNSLGFILADAGFDVWNGSRGN 135
DB 61 YRIPYGVKNSEINICKRPVAYLQHLGLIASATNITLNPNSLAFILADAGYDVLGNSRGN 120
QY 136 AWSRKHHTLSIDQDEFWAFSDENARFDPVAVINFILOKTOEKIYVYSGTGMFGIA 195
DB 121 TWSRKNVYSPDSVEFWAFSDENAKYDLPATIDFIVOKTQEKIHYVGHSGQTIGFIA 180
QY 196 FSTMPDLAOKTKMYFALAPIATVHKAKSPGKFLLLPDMIKGLFGKKEFL---YOTREFL 252
DB 181 FSTNPALAKKIKRYFALAPVATVYTESPFKISLIPFKLVIFGNKMFPHNYLDQFL 240
QY 253 RQVLYLGGVILQDQICSNIMLLGGFNTNNMNSRASVYAAHTLAGTSVQNILHWSQAV 312
DB 241 GTEV---CSRELLDLCNALFIFCGFDKKNLVSRFDVYXGHNPACTSTODLFHQAQLA 297
QY 313 NSGELRAFDMGSETKNLEKCNQPTPVRYVRDMVTVPYTAAMTGGQDWLSNPEDVKMLLSEV 372
DB 298 KSGKLQAYNWGSPQLQNMHLHYNQKTPPYVDVSANTVPIAVWNGGHDILADPDQVAMLLPKL 357
QY 373 TNLVYHKNIPEWAHVDFIWLGLDAPHMYNEIHLMOQE 410
DB 358 PNLVYHKELIPYNLHDFIWMADPAQOEYNEIVTMAED 395

RESULT 15
Q9D760 PRELIMINARY; PRT; 395 AA.
AC Q9D760
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE 2310051B21Rik protein.
GN 2310051B21Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TONGUE;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
```

```
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikola I., Pesole G., Quackenbush J.,
RA Schirai L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayaehizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK009560; BAB26359.1; -
DR MGI; MGI:1914967; 2310051B21Rik.
DR InterPro; IPR000073; Abhydrolase.
DR InterPro; IPR000734; Lipase.
DR InterPro; IPR000379; Ser_estrs_site.
DR Pfam; PF00561; abhydrolase_1.
DR PROSITE; PS00120; LIPASE_SER; 1.
SQ SEQUENCE 395 AA; 44685 MW; PEF96B65EA670BEE CRC64;

Query Match 50.08; Score 1124.5; DB 11; Length 395;
Best Local Similarity 53.3%; Pred. No. 1.7e-90;
Matches 212; Conservative 70; Mismatches 109; Indels 7; Gaps 3;

QY 17 MWLLILVAYMFO-RNVSVMHPTKAVDPPEAFMNISEIIIOHOGYPCPEEVEVATEDGYILSV 75
DB 1 MWLLIVTSVLSAFSGAGHGLFGKLPKPNPEANNVSMITYWGYPSEYEVVTEGYILGV 60
QY 76 NRIPRGLVQPKTKTSRPPVYLQHLGVGGASNMISLNPNSLGFILADAGFDVWNGSRGN 135
DB 61 YRIPYGVKNSEINICKRPVAYLQHLGLIASATNITLNPNSLAFILADAGYDVLGNSRGN 120
QY 136 AWSRKHHTLSIDQDEFWAFSDENARFDPVAVINFILOKTOEKIYVYSGTGMFGIA 195
DB 121 TWSRKNVYSPDSVEFWAFSDENAKYDLPATIDFIVOKTQEKIHYVGHSGQTIGFIA 180
QY 196 FSTMPDLAOKTKMYFALAPIATVHKAKSPGKFLLLPDMIKGLFGKKEFL---YOTREFL 252
DB 181 FSTNPALAKKIKRYFALAPVATVYTESPFKISLIPFKLVIFGNKMFPHNYLDQFL 240
QY 253 RQVLYLGGVILQDQICSNIMLLGGFNTNNMNSRASVYAAHTLAGTSVQNILHWSQAV 312
DB 241 GTEV---CSRELLDLCNALFIFCGFDKKNLVSRFDVYXGHNPACTSTODLFHQAQLA 297
QY 313 NSGELRAFDMGSETKNLEKCNQPTPVRYVRDMVTVPYTAAMTGGQDWLSNPEDVKMLLSEV 372
DB 298 KSGKLQAYNWGSPQLQNMHLHYNQKTPPYVDVSANTVPIAVWNGGHDILADPDQVAMLLPKL 357
QY 373 TNLVYHKNIPEWAHVDFIWLGLDAPHMYNEIHLMOQE 410
DB 358 PNLVYHKELIPYNLHDFIWMADPAQOEYNEIVTMAED 395

Search completed: February 3, 2003, 13:49:58
Job time : 37 secs
```